U.PORTO

MESTRADO EM ONCOLOGIA

ESPECIALIZAÇÃO EM ONCOLOGIA MOLECULAR

DNA Methylation Profiling as a Tool for Testicular Germ Cell Tumors Subtyping

Ana Laura da Silva Costa

M

2017

U.PORTO

Ana Laura Costa. DNA Methylation Profiling for Testicular Germ Cell Tumors Subtyping **M**.ICBAS **2017**

DNA Methylation Profiling as a Tool for Testicular Germ Cell Tumors Subtyping

Ana Laura da Silva Costa

INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR

ANA LAURA DA SILVA COSTA

DNA Methylation Profiling as a Tool for Testicular Germ Cell Tumors Subtyping

Dissertação de candidatura ao grau de **Mestre em Oncologia – Especialização em Oncologia Molecular** submetida ao Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto

Orientador – Professor Doutor Rui Manuel Ferreira Henrique
Professor Catedrático Convidado
Departamento de Patologia e Imunologia Molecular
Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto
Diretor do Serviço de Anatomia Patológica
Investigador Sénior do Grupo de Epigenética e Biologia do Cancro
Centro de Investigação
Instituto Português de Oncologia do Porto

Co-orientadora – Professora Doutora Carmen de Lurdes Fonseca Jerónimo
Professora Associada Convidada com Agregação
Departamento de Patologia e Imunologia Molecular
Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto
Investigadora Auxiliar e Coordenadora do Grupo de Epigenética e Biologia do Cancro
Centro de Investigação
Instituto Português de Oncologia do Porto

“It is hard to fail, but it is worse
never to have tried to
succeed”

Theodore Roosevelt

AGRADECIMENTOS

Chega ao fim mais uma etapa da minha vida, a concretização de mais um objetivo pessoal que tracei há dois anos atrás e o qual só se tornou possível de realizar devido às várias pessoas que me acompanharam neste percurso. É a todas elas que segue o meu agradecimento.

Em primeiro lugar gostaria de agradecer ao meu orientador, Professor Doutor Rui Henrique e à minha co-orientadora, Professora Doutora Carmen Jerónimo por me terem dado a oportunidade de fazer parte do Grupo de Epigenética e Biologia do Cancro. Muito obrigada por todos os ensinamentos, por todo o tempo despendido em esclarecimento de dúvidas, pela enorme disponibilidade, críticas, conselhos, paciência, apoio e por contribuírem para o meu desenvolvimento pessoal e profissional na área de investigação em cancro. Muito obrigada pela confiança.

Ao Professor Doutor Manuel Teixeira, na qualidade de Diretor do Departamento de Genética e do Centro de Investigação do Instituto Português de Oncologia (IPO) do Porto, por ter autorizado o desenvolvimento deste projeto no Centro de Investigação.

À Professora Doutora Carmen Jerónimo, na qualidade de Diretora do Mestrado em Oncologia e à sua antecessora, Professora Doutora Berta Martins, por me terem aceite no referido Mestrado.

Ao Serviço de Anatomia Patológica do IPO do Porto em particular ao Dr. João Lobo por toda a dedicação para com este projeto. Foi uma grande dor de cabeça lidar com a complexidade associada a estes tumores e sem dúvida que sem a tua sabedoria juntamente com a tua persistência, resiliência e vontade tal não seria possível. Obrigada por acreditares neste projeto e por teres trabalhado tão arduamente a vários níveis para a sua concretização. Obrigada também por todo o apoio que me deste e por todos os esclarecimentos, sobretudo no início em que tive grandes dificuldades na compreensão destes tumores. Agradeço também às Técnicas Mariana Cantante e Rita Guimarães deste mesmo serviço por toda a dedicação e prontidão na realização da enorme quantidade de cortes histológicos que lhes pedimos. Agradeço também à Técnica Paula Lopes pela realização de toda a imunohistoquímica, sem a qual não seria possível determinar parte dos componentes nos tumores mistos.

Ao Serviço de Epidemiologia do IPO do Porto na pessoa do Engenheiro Luís Antunes pelo esclarecimento de algumas dúvidas relativas à análise estatística associada a este estudo.

À Catarina Barbosa, a minha "Katezinha", obrigada por todo o conhecimento experimental e teórico que me transmitiste e continuas a transmitir, foi contigo que aprendi praticamente tudo o que neste momento sei fazer de parte experimental, foste tu que desde

início sempre mostraste uma enorme disponibilidade e vontade de ajudar sempre em tudo e para com todas as pessoas. Foste tu que pegaste em mim desamparada por nunca ter trabalhado num laboratório antes e me ensinaste da melhor forma possível como o fazer e bem. Foste tu que aos poucos te foste integrando no meu projeto, tornando-te na minha "chefinha", ajudando-me imenso a todos os níveis sobretudo nos dias em que mais "bati com a cabeça nas paredes" e quando nada me fazia sentido. Foste tu que sempre estiveste lá nas horas de maior angústia, sempre pronta a ver o lado positivo, a tentar encontrar alguma explicação para os problemas, a nunca me deixar desistir. Por tudo isto e muito mais um enorme e profundo obrigada! Obrigada por todas as horas que perdeste comigo e te dedicaste ao meu/nosso trabalho, obrigada por nunca me teres deixado sozinha e teres trabalhado comigo até tarde, pelos fins de semana que ficaste para me ajudar, por tudo o que tens feito por mim! Obrigada pela tua amizade e por toda a paciência que tiveste e tens comigo nos dias em que estive com os "azeites", obrigada por saberes lidar com a minha personalidade e por me arrancares sempre um sorriso simplesmente pela pessoa que és. Obrigada por me fazeres acreditar que ainda existem pessoas generosas, sinceras e boas em geral no mundo! Levo-te para a vida!

Às minhas compinchas do Mestrado, do Laboratório e da Vida: Bárbara Salgueiro, a minha parceira na "testis team", obrigada por toda a ajuda no trabalho prático que fizemos em conjunto, obrigada por me fazeres chorar a rir praticamente todos os dias com as tuas histórias e com simplesmente a tua maneira de ser. Tornaste-te muito mais do que uma parceira de trabalho, tornaste-te numa grande amiga, sempre pronta a ouvir os meus desesperos e agoiros e a partilhar alegrias também e por isso um enorme obrigada por teres estado sempre lá; à Daniela Silva, um grande obrigada por toda a ajuda não só a nível prático, mas também por sempre teres estado pronta a pensar e a resolver as minhas dúvidas/problemas. Obrigada por todo o ânimo e por toda a calma que sempre me tentaste transmitir em horas mais "stressantes". Obrigada por teres tornado os meus dias de processamento, as minhas horas de almoço etc, bastante mais divertidos, bem como as viagens para casa. Obrigada pela pessoa divertida e genuína que és!; à Ângela Magalhães, uma das minhas coleguinhas de casa e amiga de longa data, companheira nesta aventura que tem sido a vida universitária, obrigada por teres estado sempre lá, por continuares a estar, por todo o apoio incondicional e amizade. Obrigada por todas as vezes que cozinhas-te para mim quando cheguei tarde, obrigada por toda a companhia que me fizeste, por partilhares comigo todos os momentos da minha vida, desde choro às maiores gargalhadas, desde trabalho até diversão. Este percurso não teria sido o mesmo sem a tua presença!; à Maria Amorim, obrigada por todas as vezes que me ajudaste e me ouviste quando tinha dúvidas, e pela excelente coleguinha de casa e amiga em que te tornas-te e

desculpa ter sido uma preguiçosa e não ter acompanhado as tuas idas ao ginásio, a intenção era boa.

Aos restantes membros do GEBC: Inês, Sara, Sofia, Vera, David, Lameirinhas, Helena, Sandra, Ana Luísa e Catarina agradeço todo o apoio, disponibilidade e ajuda, sobretudo das mais velhas que sempre se mostraram disponíveis para ajudar e para transmitir os seus conhecimentos e experiências. Lameirinhas, obrigada pelas tuas sempre pertinentes palavras de tranquilidade e pela confiança que me tentaste transmitir (a célebre frase, "Tu consegues!"), mantém sempre essa tua maneira de ser tão característica e obrigada pelos choros de riso que me proporcionaste com algumas das tuas "saídas". Lameirinhas, Helena e Sandra, muito obrigada pelo tempo que dedicaram à leitura da minha tese e pelo trabalho que tiveram nas correções. Às mais novinhas desejo as maiores felicidades e sorte para os próximos tempos.

À Dona Luísa, Dona Marta e Dona Elisabete, obrigada pelas vossas palavras, carinho e sorrisos todos os dias. Dona Luísa e Dona Marta, obrigada pelo vosso trabalho para que nada nos faltasse no laboratório. Dona Elisabete, obrigada por toda a sua simpatia e boa disposição em todos os finais de tarde e pelos deliciosos pastéis de nata que fez questão de nos trazer.

Às demais pessoas, colegas e amigos de Licenciatura e da "terrinha", alguns dos quais apesar da minha enorme ausência para com eles, sempre estiveram presentes para me apoiar e nunca se esqueceram de mim.

Ao Bruno Moreira, por todo o carinho, apoio, incentivo e ajuda, especialmente durante este último ano. Acima de tudo, obrigada por sempre teres acreditado em mim, por sempre me tentares transmitir calma e confiança, mesmo estando a passar pelo mesmo. Um enorme obrigada pela grande paciência que tiveste e tens tido para comigo em dias mais difíceis em que sei que não foi fácil lidar comigo, obrigada por me ouvires, por me fazeres companhia, por sempre me tentares alegrar, por sempre me arrancares um sorriso a qualquer hora e em qualquer altura. Muito obrigada por tudo!

Por último mas ainda mais importante, agradeço aos meus pais, ao meu irmão, à minha cunhada e ao meu afilhado, por serem as pessoas que sempre me incentivaram e estiveram do meu lado e sem as quais nada do que concretizei até hoje seria possível. Mãe, Pai, obrigada pelo enorme esforço que sempre fizeram e continuam a fazer para que eu possa continuar a sonhar e a lutar pelos meus objetivos e sonhos. Obrigada por nunca me dizerem "não" a qualquer objetivo que tenha traçado por mais que isso implicasse terem de trabalhar mais ou terem de abdicar de coisas vossas para que eu o pudesse concretizar. Obrigada por todas as palavras de conforto que fizeram questão de me transmitir a cada chamada, pela paciência, compreensão e preocupação. Mano, obrigada pelo segundo pai que tens sido para mim, pelo apoio incondicional que sempre me deste, por estares sempre

disponível para qualquer coisa que precise. Obrigada pela proteção, carinho e por seres o meu ídolo e me inspirares todos os dias pelo grande Homem que és. Sandra, obrigada pelo apoio que sempre me deste e por sempre permitires que invada a vossa casa seja para trabalhar ou não. Obrigada por me tratares sempre bem, pela preocupação, pela tua disponibilidade constante e por seres uma irmã mais velha para mim. Gustavo, ainda não tens noção de nada da vida adulta nem dos problemas, mas ainda bem, essa linda criança traquina que és, sempre com um sorriso na cara tem sido uma das coisas mais reconfortantes na minha vida, mesmo sem saberes ainda. Pai, Mãe, Mano, Cunhada e Afilhado, agradeço-vos do fundo do coração por tudo, devo-vos tudo! São o melhor que podia pedir!

Esta dissertação é dedicada aos meus Pais, ao meu Irmão, à minha Cunhada e ao meu Afilhado!

This study was funded by a grant of the Research Center of Portuguese Oncology Institute of Porto (CI-IPOP 27)

RESUMO

INTRODUÇÃO: Os tumores de células germinativas testiculares (TCGT) constituem aproximadamente 1% de todos os tumores malignos no sexo masculino. No entanto, representam a neoplasia mais comum nos homens em idade jovem, nos países mais desenvolvidos. Os TCGT dividem-se em dois grupos: os relacionados com a neoplasia de células germinativas *in situ* (NCGIS) e os não relacionados. A grande maioria dos TCGT é representada pelo grupo de tumores relacionados com a NCGIS, os quais incluem dois subtipos principais, nomeadamente os seminomas (SE) e os tumores não seminomatosos (NS). Os tumores NS incluem ainda quatro subtipos principais: carcinoma embrionário (CE), tumor do saco vitelino (SV) do tipo pós-pubertal, coriocarcinoma (CC) e teratoma (TE) do tipo pós-pubertal. Todos estes subtipos de tumores NS podem surgir combinados entre si e mesmo em combinação com SE, constituindo um grupo referido como tumores de células germinativas mistos. Como se torna evidente, os TCGT são um grupo de tumores heterogéneo tanto a nível patológico como a nível clínico, apesar da sua grande maioria ter origem na mesma via de tumorigénese e partilhar a mesma alteração genética básica. Assim, a descoberta de novos biomarcadores baseados na alteração do estado de metilação no promotor de genes pode ser promissora na discriminação entre os diferentes subtipos de TCGT, incrementando a acuidade do diagnóstico e auxiliando na determinação do prognóstico e na monitorização dos doentes.

OBJETIVOS: O principal objetivo deste estudo foi avaliar o perfil de metilação de um painel de cinco genes - *CRIPTO*, *HOXA9*, *MGMT*, *RASSF1A* e *SCGB3A1* – nos principais subtipos de TCGT. Especificamente, pretendeu-se avaliar se o padrão de metilação destes genes estava associado com os diferentes subtipos de TCGT. Adicionalmente, foi avaliada a associação entre a metilação do promotor destes cinco genes e os parâmetros clinico-patológicos, bem como o potencial valor de prognóstico.

MATERIAIS E MÉTODOS: Os níveis de metilação dos promotores dos genes *CRIPTO*, *HOXA9*, *MGMT*, *RASSF1A* e *SCGB3A1* foram avaliados através de PCR (*Polymerase chain reaction*) específico de metilação, realizado em ADN genómico após modificação por bissulfito de sódio. Este ADN foi obtido a partir de amostras de tecido fixadas em formol e incluídas em parafina, provenientes de uma coorte de 161 doentes com TCGT e 16 amostras controlo. O valor de diagnóstico e de discriminação destes cinco genes foi determinado através da análise de curvas ROC (*Receiver operating characteristics*). Avaliaram-se, também, possíveis associações entre os níveis de metilação do promotor destes cinco genes e os parâmetros clinico-patológicos.

RESULTADOS: Os níveis de metilação do *CRIPTO*, *HOXA9*, *MGMT* e *SCGB3A1* foram significativamente superiores nos TCGT relacionados com a NCGIN do que nas amostras controlo. Particularmente, o painel constituído por *CRIPTO*/*HOXA9*/*SCGB3A1* apresentou um elevado desempenho de diagnóstico para tumores NS com todos os parâmetros acima de 90%, exceto especificidade e valor preditivo negativo. Os níveis de metilação do *RASSF1A* distinguiram as amostras de SE puros das amostras controlo com todos os parâmetros de desempenho de diagnóstico acima de 73%. Todos os cinco genes candidatos discriminaram significativamente SE puros de tumores NS, com níveis de metilação mais elevados observados nas amostras de tumores NS. Definiu-se, assim, um painel de genes constituído por *HOXA9* e *RASSF1A* para discriminação entre SE puros e tumores NS, cuja sensibilidade, especificidade, precisão e valor de área sob a curva foram de 90,1%, 81,6%, 87,4% e 0,904, respetivamente. Adicionalmente, observaram-se diferenças significativas nos níveis de metilação do *CRIPTO*, *MGMT* e *RASSF1A* quando comparados os SE puros com os SE como componentes dos tumores de células germinativas mistos, apenas do *MGMT* entre os CEs puros e as correspondentes formas mistas e do *HOXA9*, *RASSF1A* e *SCGB3A1* entre os TEs pós-pubertais puros e os TEs pós-pubertais provenientes de tumores mistos. Observou-se, também, que os níveis de metilação do *HOXA9*, *RASSF1A* e *SCGB3A1* foram significativamente superiores nos CEs puros quando comparados com os SE puros e, simultaneamente, que a metilação do promotor do *CRIPTO* foi significativamente superior nos TEs pós-pubertais puros quando comparados com os SE puros e com os CEs puros, individualmente. Não obstante, os níveis de metilação do *HOXA9* e do *RASSF1A* foram significativamente mais baixos nos SE mistos quando individualmente comparados com os restantes subtipos de TCGT em contexto de tumores mistos. Os níveis de metilação do *RASSF1A* foram significativamente superiores nos tumores do SV pós-pubertais e nos CCs ambos em contexto de TCGT mistos quando comparados com os CEs mistos. Os níveis de metilação do *SCGB3A1* foram significativamente superiores nos tumores do SV pós-pubertais, CCs e TEs pós-pubertais em contexto de TCGT mistos quando individualmente comparados com os SE mistos ou com os CEs mistos. Os níveis de metilação do *MGMT* foram apenas significativamente diferentes entre os SE mistos e os TEs mistos do tipo pós-pubertal, com os níveis mais elevados observados neste último componente.

Os níveis de metilação do *CRIPTO* e *HOXA9* foram também significativamente diferentes entre os tumores do SV pré-pubertais e os tumores do SV pós-pubertais como componentes dos TCGT mistos. Observaram-se, igualmente, diferenças significativas nos níveis de metilação do *HOXA9*, *MGMT* e *SCGB3A1* entre os TEs pré-pubertais e os TEs pós-pubertais em contexto de tumores mistos.

A sensibilidade dos níveis de metilação do *CRIPTO* na deteção de TCGT relacionados com a NCGIS foi superior à descrita para cada um dos marcadores tumorais séricos medidos antes da orquidectomia (sensibilidade: 65,8% *vs*. 35% de alfa-fetoproteína, 51,8% de beta-gonadotropina coriónica humana e 42,9% de desidrogenase láctica). Adicionalmente, observaram-se níveis mais elevados de metilação do *HOXA9*, *RASSF1A* e *SCGB3A1* em doentes em estadio mais avançado, particularmente em doentes em estadio III quando comparados com doentes em estadio I ou estadio II. Os níveis de metilação destes três genes associaram-se também com o estadio baseado na estratificação em grupos de prognóstico dos doentes com doença metastática, com os níveis mais elevados observados nos doentes dos grupos de mau ou prognóstico intermédio, comparado com os doentes do grupo de bom prognóstico.

CONCLUSÕES E PERSPETIVAS FUTURAS: Este estudo mostrou que a utilização de biomarcadores baseados na metilação para a deteção de TCGT é promissora, com sensibilidade superior à dos marcadores tumorais séricos clássicos. Foram, também, observadas importantes diferenças moleculares a nível epigenético entre todos os subtipos de TCGT, contribuindo para uma melhor compreensão destes tumores tão heterogéneos. Adicionalmente, os níveis de metilação destes cinco genes poderão ser úteis para melhorar as ferramentas de estratificação de risco clínico, em termos de estadio e prognóstico. Contudo, são necessários estudos adicionais, englobando maiores coortes de doentes, bem como dadores saudáveis, para confirmação da utilidade destes biomarcadores na prática clínica. A validação destes resultados em biópsias líquidas é também necessária para determinar o potencial de um novo método minimamente invasivo na deteção e discriminação dos subtipos de TCGT, com impacto na monitorização dos doentes.

ABSTRACT

BACKGROUND: Testicular germ cell tumors (TGCT) account for approximately 1% of all male malignant tumors, but represent the most common neoplasm in young Caucasian men in developed countries. TGCT can be divided into TGCT related to germ cell neoplasia *in situ* (GCNIS) and unrelated ones. The former represents the vast majority of TGCT and include seminoma (SE) and nonseminomatous tumors (NST), being the last further divided into embryonal carcinoma (EmbrCa), postpubertal-type yolk-sac tumor (YST), choriocarcinoma (CH) and postpubertal-type teratoma (TE). Different combinations of NST components and even SE may occur in a class of tumors referred to as mixed germ cell tumors. TGCT are a heterogeneous group of tumors at pathological and clinical levels, although most originate from the same tumorigenic pathway and are characterized by a common genetic abnormality. Thus, the discovery of new methylation-based biomarkers may show promise for discrimination between TGCT subtypes, providing additional relevant information at diagnosis to support therapeutic decisions and patient monitoring.

AIMS: The main objective of this study was to evaluate a promoter methylation profile based on a five-gene panel constituted by *CRIPTO*, *HOXA9*, *MGMT*, *RASSF1A* and *SCGB3A1* genes, in TGCT subtypes. Specifically, we aimed to assess whether methylation patterns might be associated with different TGCT subtypes. Moreover, we also evaluated the association between genes' promoter methylation and standard clinicopathological parameters, and determined their potential prognostic value.

MATERIALS AND METHODS: *CRIPTO*, *HOXA9*, *MGMT*, *RASSF1A* and *SCGB3A1* promoter methylation levels were evaluated in a cohort of 161 TGCT patients and 16 controls by quantitative methylation-specific polymerase chain reaction (qMSP), using bisulfite-modified genomic DNA, obtained from formalin-fixed paraffin-embedded tissue samples. Receiver operating characteristics (ROC) curve was performed to evaluate the diagnostic and discriminatory power of the panel. Correlation between standard clinicopathological parameters and genes' promoter methylation levels was also assessed.

RESULTS: *CRIPTO*, *HOXA9*, *MGMT* and *SCGB3A1* methylation levels were significantly higher in GCNIS-related TGCT compared to controls. Specifically, *CRIPTO*/*HOXA9*/*SCGB3A1* panel displayed high diagnostic coverage of NST with all performance parameters above of 90%, with exception of specificity and negative predictive value (NPV). *RASSF1A* methylation levels may distinguish pure SE from controls with all diagnostic parameters above 73%. All five candidate genes were able to significantly discriminate between pure SE and NST with the higher methylation levels depicted in NST. *HOXA9*/*RASSF1A* panel showed high discriminatory performance for pure SE *vs.* NST,

with 90.1% sensitivity, 81.6% specificity, 87.4% accuracy and area under the curve (AUC) of 0.904. Moreover, significant differences for *CRIPTO*, *MGMT* and *RASSF1A* methylation levels were depicted comparing pure SE with SE as component of mixed TGCT, for *MGMT* between pure EmbrCa and matching mixed EmbrCa and for *HOXA9*, *RASSF1A* and *SCGB3A1* between pure postpubertal-type TE and mixed postpubertal-type TE. *HOXA9*, *RASSF1A* and *SCGB3A1* methylation levels were also significantly higher in pure EmbrCa compared with pure SE, whereas *CRIPTO* promoter methylation levels were significantly higher in pure postpubertal-type TE compared to pure SE and pure EmbrCa. Moreover, *HOXA9* and *RASSF1A* methylation levels were significantly lower in mixed SE when individually compared with the remaining mixed TGCT subtypes. Furthermore, *RASSF1A* methylation levels were also significantly higher in mixed postpubertal-type YST and mixed CH when compared with mixed EmbrCa. *SCGB3A1* methylation levels were significantly higher in mixed postpubertal-type YST, mixed CH and mixed postpubertal-type TE when individually compared with mixed SE or mixed EmbrCa. *MGMT* methylation levels were only significantly different between mixed SE and mixed postpubertal-type TE with the highest levels depicted in the last component. A significant difference was also observed for *CRIPTO* and *HOXA9* methylation levels between prepubertal-type YST and postpubertal-type YST as component of mixed germ cell tumors and for *HOXA9*, *MGMT* and *SCGB3A1* methylation levels between prepubertal-type TE and mixed postpubertal-type TE.
Diagnostic information on GCNIS-related TGCT was improved when considering *CRIPTO* methylation levels than the three classical serum tumor markers measured before orchiectomy (sensitivity: 65.8% *vs*. 35.0% of alpha-fetoprotein, 51.8% of beta-human chorionic gonadotropin and 42.9% of lactate dehydrogenase). Additionally, patients with advanced group staging disclosed the highest methylation levels of *HOXA9*, *RASSF1A* and *SCGB3A1*, with the highest methylation levels depicted in patients with stage III disease compared with patients at stage I or stage II. Promoter methylation levels of these genes panel were also significantly associated with prognostic-based staging of TGCT with the higher methylation levels depicted by patients classified as intermediate and poor prognosis compared with good prognosis patients.

CONCLUSIONS AND FUTURE PERSPECTIVES: This study illustrated the promising use of methylation-based biomarkers for TGCT detection, with higher sensitivity than the standard classical serological markers. We also observed molecular differences at epigenetic level between all TGCT subtypes, which might aid in management of these heterogeneous tumors. Moreover, we also showed the potential value of combinations of these five genes promoter methylation levels for improving risk stratification and prognostication of TGCT patients. However, further studies using larger cohorts of patients

and healthy donors are mandatory to confirm the usefulness of these promising biomarkers in clinical practice. Validation of these results in liquid biopsies might also provide a novel minimally invasive tool for detection and discrimination among TGCT subtypes, as well as patient monitoring.

TABLE OF CONTENTS

FIGURE INDEX

TABLE INDEX

LIST OF ABREVIATIONS

µL	Microliter
µm	Micrometer
5mC	5-methylcytosine
A	Adenine
AFP	Alpha-fetoprotein
AP-2γ	Transcription factor activator protein 2γ
APC	Adenomatous polyposis coli
AUC	Area under the curve
AZFc	Azoospermia factor c
BAK1	BCL2 antagonist/killer 1
BEP	Cisplatin, etoposide, bleomycin
bp	Base pairs
C	Cytosine
CBEG	Cancer Biology and Epigenetic Group
CH	Choriocarcinoma
CH_3	Methyl group
cm	Centimeter
CpG	Cytosine-phosphate-Guanine
CRIPTO (TDGF1)	Teratocarcinoma-derived growth factor 1
CT	Abdominopelvic computed tomography
DMRT1	Doublesex and mab-3 related transcription factor 1
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
dsDNA	Double-strand DNA
EmbrCa	Embryonal carcinoma
EP	Etoposide, cisplatin
ER-β	Estrogen receptor beta
ESC	Embryonic stem cell
F	Forward
FFPE	Formalin-fixed paraffin-embedded
FGFR3	Fibroblast growth factor receptor 3
FKBP4	FK506 binding protein 4
G	Guanine
GATA3	GATA binding protein 3
GCNIS	Germ cell neoplasia *in situ*

H&E	Hematoxylin and Eosin
HOXA9	Homeobox A9
HRAS	Hras proto-oncogene, GTPase
i12p	Isochromosome 12p
IGCCCG	International Germ Cell Cancer Collaborative Group
IGD	In greatest dimension
IQR	Interquartile range
IU/L	International unit/Liter
KIT	KIT proto-oncogene receptor tyrosine kinase
KITLG	KIT ligand
KRAS	KRAS proto-oncogene GTPase
LDH	Lactate dehydrogenase
lncRNA	Long non-coding RNA
MAGEA4	Melanoma antigen family member A4
MBD	Methyl-CpG-binding
MeCP2	Methylcytosine-binding protein 2
mg	Milligram
MGMT	O-6-methylguanine-DNA methyltransferase
miRNA	microRNA
mL	Milliliter
MLH1	MutL homolog 1
mRNA	Messenger RNA
NANOG	Nanog homeobox
ncRNA	Non-coding RNA
ng	Nanogram
nM	Nanomolar
NPV	Negative predictive value
NPVM	Nonpulmonary visceral metastases
NST	Nonseminomatous tumors
ºC	Celsius degree
PCR	Polymerase chain reaction
PDE11A	Phosphodiesterase 11A
PFS	Progression free survival
PGC	Primordial germ cell
PLAP	Placental alkaline phosphatase
POU5F1 **(*OCT3/4*)**	Pou class 5 homeobox 1

PPV	Positive predictive value
pTNM	Pathological TNM classification
qMSP	Quantitative methylation-specific polymerase chain reaction
R	Reverse
RASSF1A	Ras association domain family member 1 isoform A
RNA	Ribonucleic acid
ROC	Receiver operating characteristics
RPLND	Retroperitoneal lymph node dissection
SAH	S-adenosyl-homocysteine
SALL4	Spalt like transcription factor 4
SAM	S-adenosyl-methionine
SCGB3A1	Secretoglobin family 3A member 1
SE	Seminomas
sncRNA	Small non-coding RNA
SOX17	SRY-box 17
SOX2	SRY-box 2
SPRY4	Sprouty RTK signaling antagonist 4
ST	Spermatocytic tumor
T	Thymine
TE	Teratoma
TGCT	Testicular germ cell tumors
TNM	Tumor Node Metastasis
TSG	Tumor suppressor genes
U	Uracil
UK	United Kingdom
ULN	Upper limit of normal range
UTR	Untranslated region
VIP	Etoposide, cisplatin, ifosfamide
WHO	World Health Organization
XIST	X inactive specific transcript (non-protein coding)
YST	Yolk-sac tumor
β-hCG	Beta-human chorionic gonadotropin

I. INTRODUCTION

1. TESTICULAR GERM CELL TUMORS

Testicular germ cell tumors (TGCT) are a very complex and heterogeneous group of neoplasms characterized by diagnostic, therapeutic and prognostic particularities. Currently, according to the 2016 World Health Organization (WHO) classification, TGCT are divided into two main categories: those related to germ cell neoplasia *in situ* (GCNIS) and the unrelated ones [1]. Importantly, since TGCT unrelated to GCNIS are extremely rare, the global epidemiological data as well as the risk factors reported in the literature for TGCT only refers to TGCT related to GCNIS.

1.1. Epidemiology

TGCT are relatively rare, accounting for about 1% of all male cancers worldwide. Nevertheless, these tumors constitute the most common cancer in young Caucasian men between 15 and 44 years in developed countries [2-4]. Indeed, the highest incidence rates occur in more developed and industrialized countries, particularly in Northern and Western Europe, while the lowest rates correspond to Asia and Africa (Figure 1) [2, 4].

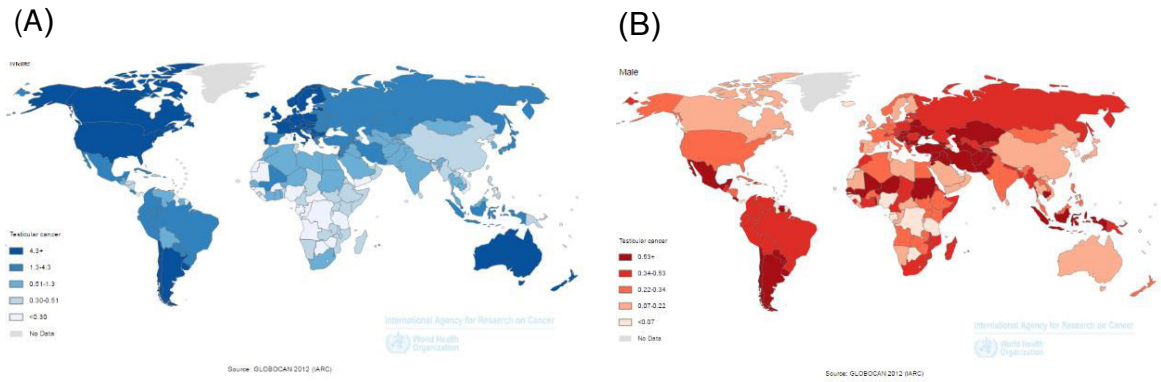


Figure 1. (A) Worldwide incidence of testicular germ cell tumors; (B) Worldwide mortality of testicular germ cell tumors. (From Globocan 2012 [3]).

1.2. Risk Factors

Currently, there are several risk factors established for TGCT development, specifically those related with genetic and environmental factors [5, 6]. About 25% of TGCT cases are attributable to genetic susceptibility which might be explained, in part, by the presence of *gr*/*gr* microdeletion (removal of part of AZFc region) at Y chromosome [7]. This molecular alteration consists in a low penetrance susceptibility allele, whose presence, possibly due to its relation with male infertility, augments in 2 to 3-fold the risk of sporadic and familial TGCT development, respectively [8-10]. Moreover, *PDE11A* gene mutations have also been described as risk modifier for familial TGCT, such as reported by Horvath *et al.,* in

which 19% of families studied displayed all *PDE11A*-gene variants [10, 11]. Additionally, United Kingdom (UK) genome-wide association study for TGCT susceptibility reported other susceptibility loci (*KITLG*, *SPRY4* and *BAK1)* associated with 6-fold risk of TGCT development for *KITLG* and 2-fold for variants in the region of *SPRY4* and *BAK1* [12]. Familial predisposition is a well-established risk factor for TGCT development representing one of the strongest known risk factors for TGCT, as proven by Hemminki *et al*., in which the overall familial risks were 3.78 and 8.78 for son-father and brothers relations, respectively [6, 8, 13, 14]. Although, other risk factors have been described, namely patients with cryptorchidism history (2 to 8-fold increased risk of TGCT development) [15, 16], subfertility/infertility (increased risk for TGCT in 1.6-2.8 times) [17-20] and testicular microlithíasis (13.2 relative risk) [21]. Furthermore, it was also described that patients with TGCT in one testis have a 25-fold increased risk to develop a contralateral tumor [15].

Somatic genetic alterations such as activating mutations in *KIT* and *KRAS* oncogenes and pathognomonic gain of chromosome 12p as isochromosome 12p (i12p) have been reported in TGCT and, therefore, may be associated with the risk of developing these neoplasms [10]. *KIT* mutation/amplification is present in about 20% of seminomas (SE) and rarely seen in nonseminomatous tumors (NST), whereas *KRAS* mutations are described in about 10% of TGCT [22-24].

On the other hand, perinatal factors such as maternal bleeding, birth order, number of siblings, inguinal hernia, twinning, low birth weight and environmental risk factors namely dyethylstillbestrol exposure *in utero*, organochloride pesticides exposure, firefighting, aircraft maintenance, leather and metal workers have also been associated with TGCT development [5, 6, 15, 25]. Maternal smoking, diet, physical exercise and marijuana intake have also been associated with increased risk of TGCT development [5, 6, 25]. It has also been reported a possible association between genvironmental factors and an increased risk of TGCT. These consist in the combination of genetic and environmental factors, resulting in the called testicular dysgenesis syndrome, which comprises, among others, cryptorchidism and testicular atrophy [5, 26, 27].

1.3. Testicular Germ Cell Tumors Subtypes

TGCT can be generally referred as GCNIS-derived and non-GCNIS-derived tumors [1]. The main subtypes of each category are grouped in Table 1 and illustrated in Figure 2.

Table 1. Classification of testicular germ cell tumors, according to WHO 2016. (Adapted from [25]).

CLASSIFICATION	TUMOR SUBTYPES
Testicular germ cell tumors related to germ cell neoplasia *in situ*	
Non-invasive germ cell neoplasia	Germ cell neoplasia *in situ* (GCNIS)
	Specific forms of intratubular germ cell neoplasia
Invasive postpubertal-type TGCT	Seminoma
	Nonseminomatous tumor: Embryonal carcinoma
	Nonseminomatous tumor: Yolk-sac tumor
	Nonseminomatous tumor: Choriocarcinoma
	Nonseminomatous tumor: Teratoma
	Nonseminomatous tumor: Mixed germ cell tumors
Testicular germ cell tumors unrelated to germ cell neoplasia *in situ*	
Prepubertal-type TGCT	Yolk-sac tumor
	Teratoma
Spermatocytic tumor	

1.3.1. Testicular Germ Cell Tumors Related to Germ Cell Neoplasia *in situ*

TGCT related to GCNIS represent the vast majority of TGCT and can be divided into SE and NST, with the last further including different subtypes according to its differentiation (Table 1) [25]. These tumors share many characteristics: all of them result from GCNIS cells progression, arise from postpubertal testis, are cytogenetically similar (all showing the gain of chromosome 12p, mainly as i12p) and display malignant behavior [1, 25, 28].

1.3.1.1. Germ Cell Neoplasia *in situ*

GCNIS consists of abnormal germ cells/gonocytes whose maturation to prespermatogonia was stopped during the normal fetal development [25, 28]. These cells remain dormant until puberty and, at this stage, GCNIS cells can progress to SE, which represents the default pathway, or to embryonal carcinoma cells whose diverse differentiation pathways originate the NST subtypes: stem-cell like embryonal carcinoma (EmbrCa), postpubertal-type yolk-sac tumor (YST) and choriocarcinoma (CH), both representing the extraembryonic lineages; and postpubertal-type teratoma (TE), representing the somatic lineage [25, 28] (Figure 2). Since GCNIS is found in almost all SE and NST, it is considered the noninvasive precursor for TGCT. About 50% of patients diagnosed with GCNIS will progress to an invasive TGCT within 5 years, whereas 70% will

progress to an invasive TGCT in 7 years, thus, suggesting that almost all GCNIS progress to overt cancer [27, 29]. Furthermore, GCNIS is found in about 5% of contralateral testis of previous TGCT patients [30].

GCNIS cells express various pluripotency markers such as POU5F1 (OCT3/4), placental alkaline phosphatase (PLAP), AP-2γ, NANOG, LIN28 and podoplanin, which might assist in diagnosis [31-35].

The highest risk to develop GCNIS has been associated with sex development disorders (70% of prevalence) and testicular dysgenesis syndrome [35, 36], being testicular biopsy the gold standard for its detection [29, 30].

Epigenetically, GCNIS cells are characterized by global DNA hypomethylation [37], presence of permissive histone modifications [38] and overexpression of miR-371-373 [39].

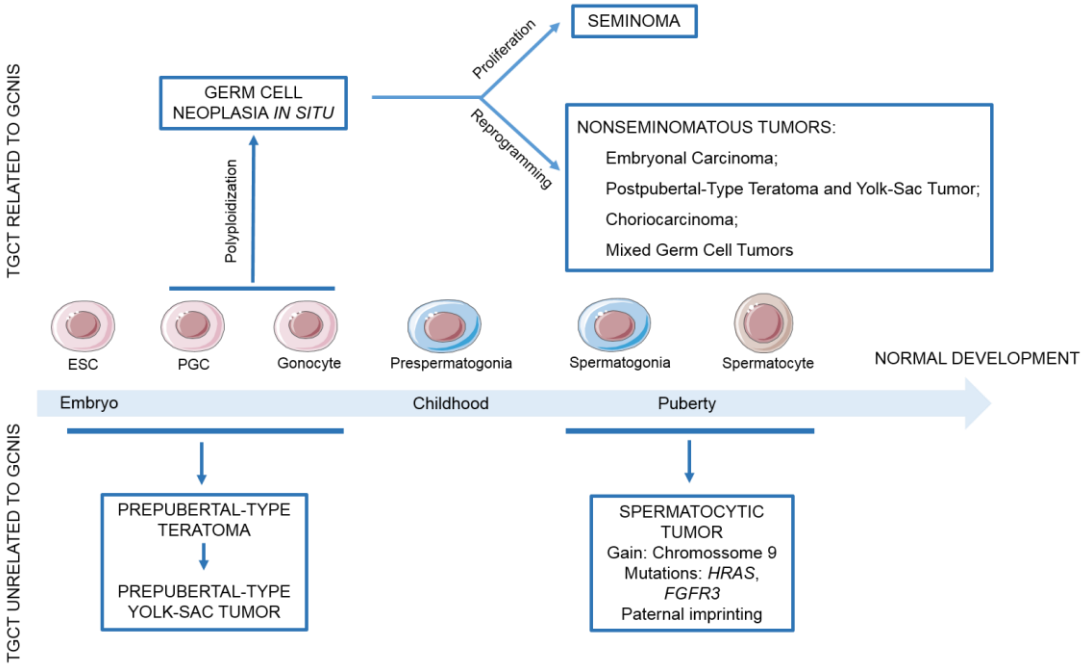


Figure 2. Pathogenesis of testicular germ cell tumors (TGCT). TGCT can be divided in two major categories: those related to germ cell neoplasia *in situ* (GCNIS) and the unrelated ones. GCNIS originates from a dormant, transformed primordial germ cell (PGC) or gonocyte, whose maturation to prespermatogonia was prevented possible due to an early polyploidization. After puberty, these transformed germ cells progress originating seminoma or one of the nonseminomatous tumors: stem-cell like embryonal carcinoma, postpubertal-type yolk-sac tumor and choriocarcinoma (extraembryonic lineages), postpubertal-type teratoma (somatic lineage) and mixed germ cell tumors composed by at least two of these different components and possible also seminoma. TGCT unrelated to GCNIS include prepubertal-type teratoma and yolk-sac tumor and spermatocytic tumor. Prepubertal-type teratoma derives from an embryonic stem cell (ESC), PGC or gonocyte limited in their development. Prepubertal-type yolk-sac tumor can result from teratoma cells progression. Spermatocytic tumor originates from spermatogonia or spermatocyte. Gain of chromosome 9 as well as activating mutations in *HRAS* and *FGFR3* genes have been associated to spermatocytic tumor carcinogenesis.

1.3.1.2. Seminoma

According to the most updated WHO classification, SE is defined as a “malignant germ cell tumor whose cells are considered the neoplastic counterparts of the primordial germ cells/gonocytes present during early embryonic development” [25].

SE constitute the most common TGCT subtype, whose pure form represents about 50% of cases, with a mean patients’ age at diagnosis of 35 years [25, 27, 28]. SE might also be presented as a component of mixed germ cell tumors together with one or more NST subtypes [40]. It occurs almost exclusively in postpubertal testis, being extremely rare in prepubertal children and in the elderly men (>70 years) [40, 41]. Morphologically, SE are similar to GCNIS, and also express PLAP, KIT, OCT3/4, SALL4, SOX17 and podoplanin. Conversely, SE cells are negative for CD30, glypican 3, SOX2, alpha-fetoprotein (AFP) and beta-human chorionic gonadotropin (β-hCG) [25, 40, 42]. Globally, the standard serum tumor markers [AFP, β-hCG and lactate dehydrogenase (LDH)] are not significantly elevated in SE tumors, especially in early stages of the disease [25].

SE development has been associated with cryptorchidism history and immunodeficiency disorders [25].

SE are usually sensitive to cisplatin-based chemotherapy and radiotherapy, and display an overall good prognosis, although stage-dependent [27].

Similarly to GCNIS, the epigenetic pattern of SE cells is characterized by DNA hypomethylation and permissive histone modifications [25].

1.3.1.3. Nonseminomatous Tumor: Embryonal Carcinoma

EmbrCa is a “malignant germ cell tumor composed of tumor cells resembling embryonic stem cells” [25]. EmbrCa represents the second most common TGCT subtype and the most common NST subtype, occurring mostly as a component of mixed germ cell tumors (about 84%) and being rare in its pure form (about 16%) [40, 43]. The mean patients’ age at diagnosis ranges between 25 to 35 years and it does not occur in prepubertal patients, being rare in men over 50 years [40].

Immunohistochemically, EmbrCa cells express OCT3/4, CD30, SOX2, SALL4, cytokeratin AE1/AE3, whereas no expression is shown for SOX17, glypican 3, KIT, podoplanin, β-hCG, carcinoembryonic antigen and epithelial membrane antigen. The expression pattern can be useful to distinguish this NST subtype from other TGCT subtypes (SE, YST or CH) [25, 40, 42, 43]. Similarly to SE, elevated serum tumor markers (AFP, β-hCG and LDH) are not usually found in pure EmbrCa form, unless EmbrCa cells co-exist with other TGCT elements [25].

Pure EmbrCa is more aggressive than EmbrCa in mixed germ cell tumors, whose prognosis depends on disease's clinical and pathological stage and the proportion of EmbrCa cells in mixed tumors [44, 45].

1.3.1.4. Nonseminomatous Tumor: Postpubertal-Type Yolk-Sac Tumor

Postpubertal-type YST is defined as "a malignant germ cell tumor that differentiates to resemble extraembryonic structures, including yolk-sac, allantois and extraembryonic mesenchyme" [25]. Postpubertal-type YST is extremely rare in its pure form and appears predominantly as a component of mixed germ cell tumors (42% of cases) [46, 47]. Patients diagnosed with these tumors usually have ages between 15-40 years, with rare cases diagnosed in the elderly [25, 48].

YST cells exhibit immunoreactivity for PLAP, α_1-antitrypsine, KIT, SALL4, pancytokeratin, AFP and glypican 3, but are negative for β-hCG, OCT3/4 and CD30 [25, 40, 49, 50]. There is a strong correlation between elevated AFP serum levels and the presence of YST, since these levels are seen in more than 95% of YST patients, thus being an important diagnostic tool [40, 42].

1.3.1.5. Nonseminomatous Tumor: Choriocarcinoma

CH is "a malignant germ cell tumor that differentiates to resemble the trophoblastic cells of the extraembryonic chorion, including cytotrophoblastic, intermediate trophoblastic and syncytiotrophoblastic cells" [25]. CH is the rarest TGCT, whose pure tumor represent less than 1% of TGCT [46, 51, 52]. Patients diagnosed with CH usually have between 20 to 39 years and often present metastatic disease, contrarily to the others TGCT subtypes. [52]. Therefore, CH is the most aggressive form of TGCT, due to the diagnosis at advanced stages, early dissemination and hemorrhagic complications [51-53]. Moreover, about 79% of patients treated with orchiectomy and chemotherapy decease of the disease within 3 years [51].

An important tool for CH diagnosis are the extremely high β-hCG serum levels, which often present levels over 50.000 IU/L [51]. Concomitantly, specific symptoms such as gynecomastia and hyperthyroidism due β-hCG similarity with luteinizing and thyroid-stimulating hormones are also common [40]. CH cells express β-hCG, SALL4, GATA3, glypican 3, α-inhibin and human placental lactogen [42].

1.3.1.6. Nonseminomatous Tumor: Postpubertal-Type Teratoma

Postpubertal-type TE is “a malignant germ cell tumor composed of several types of tissue representing one or more of the germinal layers (endoderm, mesoderm and ectoderm). It may be composed exclusively of well-differentiated, mature tissue or have immature, embryonic-type tissues” [25]. Forty seven to fifty% of postpubertal-type TE are diagnosed as a component of mixed germ cell tumors, whereas pure tumor is quite rare, occurring in only 2.7 to 7% of TGCT cases [47, 54]. This tumor mostly affects young adults and one-third are metastatic at the time of diagnosis [54]. Although, pure TE is the most common component present in patients with recurrent or metastatic disease after treatment, it associates with a more favorable prognosis compared to the presence of the other metastatic TGCT subtypes [55].

1.3.1.7. Nonseminomatous Tumor: Mixed Germ Cell Tumors

Mixed germ cell tumors are malignant tumors composed by a mixture of two or more malignant TGCT components: at diagnosis, the average patients’ age is of 30 years, being extremely rare in prepubertal patients [25, 40]. Mixed tumors that include SE component occur later than mixed tumors only composed by NST subtypes [40, 47]. EmbrCa is the most frequent subtype of mixed germ cell tumors, followed by postpubertal-type TE, postpubertal-type YST, SE and CH. All combinations are possible and usually with more than two components. The most common combinations include EmbrCa with postpubertal-type TE, SE or postpubertal-type YST [42].

Elevated serum tumor markers correlate with TGCT components present in mixed tumors. Specifically, elevated AFP levels indicates the presence of YST elements, whereas β-hCG levels are indicative of trophoblastic cells such as CH cells [40, 47]. The different constituents (SE and NST) in mixed germ cell tumors are histologically similar to their pure forms and their proportions have clinical implications, particularly in stage I tumors, thus requiring the assessment of the percentages of each component [40, 56].

1.3.2. Testicular Germ Cell Tumors Unrelated to Germ Cell Neoplasia *in situ*

TGCT unrelated to GCNIS are very rare and include spermatocytic tumor (ST) and both prepubertal-type YST and TE. Contrarily to TGCT related to GCNIS, these are a heterogeneous group with a benign behavior characterized by the absence of i12p amplification. Prepubertal-type tumors occur in children, although it may be identified in postpubertal patients. Postpubertal-type tumors may also occur in prepubertal testis [1].

1.3.2.1. Spermatocytic Tumor

ST is “a germ cell tumor derived from postpubertal-type germ cells, whose tumor cells resemble spermatogenic cells, most commonly spermatogonia or early primary spermatocytes” [25]. ST accounts only for about 1% of TGCT occurring in a wide range of ages (19 to 92 years), but with a global mean of 50 years [57]. These tumors are exclusively confined to testis and rarely metastasize, being associated with an excellent prognosis except when it progresses to sarcoma, in which a higher metastatic rate is observed [57-59]. Amplification of chromosome 9 results in *DMRT1* gene expression and *HRAS* and *FGFR3* mutations have been suggested to trigger tumor development [60, 61]. ST cells do not express embryonic germ cell markers like OCT3/4, PLAP, AFP, β-hCG, CD30, whereas proteins expressed in spermatogonia such as SALL4, KIT or MAGEA4 are often expressed [42, 62]. Epigenetic characterization of these tumors is extremely difficult since the same tumor can exhibit several levels of DNA methylation and contain permissive or repressive histone modifications [63].

1.3.2.2. Prepubertal-Type Teratoma

Prepubertal-type TE has a similar definition of postpubertal-type TE, since it is also “composed by elements resembling somatic tissues derived from one or more of the germinal layers” [25]. However, unlike postpubertal-type TE, prepubertal-type TE is a benign tumor thought to be derived from a germ cell that has not undergone malignant transformation. Moreover, these tumors particularly occur as a pure tumor and have not been associated with metastases or disease recurrence, allowing conservative treatment [42, 59, 64]. Prepubertal-type TE occurs at a median age of 13 months [64].

1.3.2.3. Prepubertal-Type Yolk-Sac Tumor

Prepubertal-type YST has the same definition of postpubertal-type YST, according to the 2016 WHO classification [25]. This is a rare tumor, although represents the most common TGCT (48% to 62% of cases) in children of 16-20 months [65, 66]. Contrarily to its postpubertal counterpart, prepubertal-type YST mostly occurs as a pure tumor with extremely rare cases reported as mixed tumors [59, 65]. Prepubertal-type YST exhibits the same morphological and immunoexpression patterns of postpubertal-type YST, being AFP serum levels elevated in about 95% of the cases [65, 67]. Nonetheless, the elevated values should be carefully interpreted, particularly in patients younger than 6 months, since they might result from the physiologically elevated values in this particular young patients [65]. About 80% of children diagnosed display stage I prepubertal-type YST with good prognosis [65].

1.4. Diagnosis and Staging

The primary diagnostic approach when there is a suspicion of testicular tumor consists in a scrotal ultrasonography of both testis, whose sensitivity in detecting testicular cancer is almost 100% [68]. The following approaches include chest and abdominopelvic computed tomography (CT), as well as measurement of classical serum tumor markers (AFP, β-hCG and LDH) for correct diagnosis and disease staging assessment [69]. Nevertheless, due to their low sensitivity and specificity, serum tumor markers should be careful interpreted. Indeed, only 60% of TGCT patients exhibit elevated serum markers at diagnosis [70]. AFP levels might also be elevated in other cancer patients, such as hepatocellular and gastrointestinal cancers, as well as in benign liver pathologies. Similarly, elevated β-hCG levels might also be associated with neuroendocrine, renal, bladder and lung cancers. Some benign and malignant diseases also result in high LDH serum levels [71, 72].

TGCT clinical staging is ascertained by clinical examination, disease extension evaluation (involvement and size of regional lymph nodes, nonregional lymph nodes and pulmonary and nonpulmonary metastases) through thorax and abdominopelvic CT and by determination of the classical serum tumor markers values before orchiectomy. Thus, extent of primary tumor (T), regional lymph nodes involvement (N), distant metastasis (M) and status of serum tumor markers (S) can be established [73].

Pathologic staging is determined after radical orchiectomy through histopathologic evaluation [73]. TGCT preferentially metastasizes through lymphatic channels to retroperitoneum rather than to haematogenous spread. In the absence of metastatic disease, serum tumor markers should be similar to basal concentrations, after orchiectomy [44, 72]. The clinical and pathologic Tumor Node Metastasis (TNM) classification of TGCT as well as the different group stages of the disease are described in Table 2 and Table 3, respectively.

Table 2. Clinical and pathological TNM classification of testicular germ cell tumors. (Adapted from [25, 69, 73]).

Clinical TNM Classification	
T - Primary tumor	
Except for pTis and pT4, for which radical orchiectomy is not always necessary for classification purposes, the extent of the primary tumor is classified after radical orchiectomy (pathologic stage). Otherwise, TX is used if radical orchiectomy has not been performed.	
N - Regional lymph nodes	
NX	Regional lymph nodes cannot be assessed
N0	No regional lymph node metastasis
N1	Metastasis with a lymph node mass ≤ 2 cm IGD or multiple lymph nodes, none > 2 cm IGD
N2	Metastasis with a lymph node mass >2 cm but not >5 cm, or multiple lymph nodes with any one mass >2 cm but not >5cm IGD
N3	Metastasis with a lymph node mass >5 cm IGD
M - Distant metastasis	
MX	Distant metastasis cannot be assessed
M0	No distant metastasis
M1	Distant metastasis
M1a	Nonregional lymph nodes or lung
M1b	Other sites
Pathological TNM classification (pTNM)	
T – Primary tumor	
pTX	Primary tumor cannot be assessed
pT0	No evidence of primary tumor (e.g histological scar in testis)
pTis	Germ cell neoplasia *in situ*
pT1	Tumor limited to testis and epididymis without vascular/lymphatic invasion; tumor may invade tunica albuginea, but not tunica vaginalis
pT2	Tumor limited to testis and epididymis with vascular/lymphatic invasion or tumor extending through into the tunica albuginea with involvement of tunica vaginalis
pT3	Tumor invades spermatic cord with or without vascular/lymphatic invasion
pT4	Tumor invades scrotum with or without vascular/lymphatic invasion
pN - Regional lymph nodes	
pNX	Regional lymph nodes cannot be assessed
pN0	No regional node metastasis
pN1	Metastasis with a lymph node mass ≤2 cm IGD and ≤5 positive nodes, none >2 cm IGD
pN2	Metastasis with a lymph node mass >2 cm but not >5 cm IGD; or >5 nodes positive, none >5 cm IGD; or evidence of extranodal tumor extension
pN3	Metastasis with lymph node mass >5 cm IGD
S - Serum tumor markers	
SX	Serum markers studies not available or not performed
S0	Serum markers levels within normal limits
S1	LDH <1.5 x ULN and β-hCG <5000 mIU/mL and AFP <1000 ng/mL
S2	LDH 1.5-10 x ULN or β-hCG 5000-50 000 mIU/mL or AFP 1000-10 000 ng/mL
S3	LDH >10 x ULN or β-hCG >50 000 mIU/mL or AFP >10 000 ng/mL
IGD: In greatest dimension; LDH: Lactate dehydrogenase; β-hCG: Beta-human chorionic gonadotropin; AFP: Alpha-fetoprotein; ULN: Upper limit of normal range.	

Table 3. Staging grouping of testicular germ cell tumors. (Adapted from [25, 69, 73]).

Stage Group	T	N	M	S
Stage 0	pTis	N0	M0	S0,SX
Stage I	pT1-4	N0	M0	SX
Stage IA	pT1	N0	M0	S0
Stage IB	pT2	N0	M0	S0
	pT3	N0	M0	S0
	pT4	N0	M0	S0
Stage IS	Any pT/TX	N0	M0	S1-3 (measured post-orchiectomy)
Stage II	Any pT/TX	N1-3	M0	SX
Stage IIA	Any pT/TX	N1	M0	S0
	Any pT/TX	N1	M0	S1
Stage IIB	Any pT/TX	N2	M0	S0
	Any pT/TX	N2	M0	S1
Stage IIC	Any pT/TX	N3	M0	S0
	Any pT/TX	N3	M0	S1
Stage III	Any pT/TX	Any N	M1	SX
Stage IIIA	Any pT/TX	Any N	M1a	S0
	Any pT/TX	Any N	M1a	S1
Stage IIIB	Any pT/TX	N1-3	M0	S2
	Any pT/TX	Any N	M1a	S2
Stage IIIC	Any pT/TX	N1-3	M0	S3
	Any pT/TX	Any N	M1a	S3
	Any pT/TX	Any N	M1b	Any S

1.5. Prognosis and Therapeutic Approaches

TGCT are one of the most curable solid tumors [44, 72]. About 75% of TGCT patients present clinical stage I (80% of SE and 60% of NST) at the time of diagnosis [74]. In this stage, about 80% of SE cases and 70% of NST benefit from surgery. Moreover, stage II or III patients also present high survival rates after orchiectomy and adjuvant treatment. Indeed, the 5-year overall survival rates for localized, regional and distant disease are of 99.2%, 96% and 73.1%, respectively [75]. Furthermore, late disease relapse only occurs in 1.4% of SE and 3.2% of NST [76]. In 1997, the International Germ Cancer Collaborative Group developed a prognostic-based staging for advanced disease based on histology, location of primary tumor, metastases and serum tumor markers values immediately before chemotherapy administration, to assist in the decision of further treatment (Table 4) [69, 72, 77, 78].

Table 4. Prognostic-based staging system for metastatic germ cell cancer (International Germ Cell Cancer Collaborative Group). (Adapted from [69, 77, 78]).

Prognosis group	Type of tumor	Criteria
Good	Seminoma (90% of cases) 5-year PFS: 82% 5-year overall survival: 86%	All of the following criteria: • Any primary site • No NPVM • Normal AFP • Any β-hCG • Any LDH
	Nonseminoma (56% of cases) 5-year PFS: 89% 5-year overall survival: 92%	All of the following criteria: • Testis/retroperitoneal primary • No NPVM • AFP < 1000 ng/mL • β-hCG < 5000 IU/L (1000 ng/mL) • LDH < 1.5 x ULN
Intermediate	Seminoma (10% of cases) 5-year PFS: 67% 5-year overall survival: 72%	All of the following criteria: • Any primary site • NPVM • Normal AFP • Any β-hCG • Any LDH
	Nonseminoma (28% of cases) 5-year PFS:75% 5-year overall survival: 80%	All of the following criteria: • Testis/retroperitoneal primary • No NPVM • AFP 1000 – 10 000 ng/mL or • β-hCG 5000 – 50 000 IU/L or • LDH 1.5 - 10 x ULN
Poor	Seminoma	No patients classified as poor prognosis
	Nonseminoma (16% of cases) 5-year PFS 41% 5-year overall survival 48%	Any of the following criteria: • Mediastinal primary • NPVM • AFP > 10 000 ng/mL or • β-hCG > 50 000 IU/L (10 000 ng/mL) or • LDH > 10 x ULN
PFS: Progression free survival; NPVM: Nonpulmonary visceral metastases; AFP: Alpha-fetoprotein; LDH: Lactate dehydrogenase; β-hCG: Beta-human chorionic gonadotropin; ULN: Upper limit of normal range.		

Stage I SE patients' have a higher relapse probability if the tumor is larger than 4 cm and invades rete testis [79, 80]. Concerning NST stage I, blood or lymphatic invasion constitutes the most important metastases predictor [81].

Gold standard treatment of primary TGCT is orchiectomy, then the patients can stay in active surveillance (for example in the case of stage I SE) or can be proposed to adjuvant treatments, including chemotherapy, radiotherapy or retroperitoneal lymph node dissection, according to different stages of the disease (Table 5) [69].

Table 5. Summary of the different therapeutic schemes for treatment of testicular germ cell tumors [69].

<table>
<tr><th colspan="2">SEMINOMA</th><th colspan="2">NONSEMINOMA</th></tr>
<tr><td colspan="4">Clinical stage I</td></tr>
<tr><td colspan="2">Active surveillance</td><td colspan="2">Active surveillance</td></tr>
<tr><td colspan="2">Adjuvant chemotherapy
• 1 cycle of Carboplatin AUC x 7</td><td colspan="2">Adjuvant chemotherapy
• 1 cycle of BEP</td></tr>
<tr><td colspan="2">Adjuvant radiotherapy
(No recommended for young patients)</td><td colspan="2">Risk-adapted treatment
Low risk (no vascular invasion):
• Surveillance (standard)
• Adjuvant chemotherapy, 1 cycle BEP
• Nerve-sparing RPLND
High risk (vascular invasion)
• Adjuvant chemotherapy, 1 cycle BEP (standard)
• Nerve-sparing RPLND
• Surveillance</td></tr>
<tr><td colspan="2">Risk-adapted treatment
(Low or high risk of metastatic disease)</td><td colspan="2">RPLND</td></tr>
<tr><td colspan="4">Clinical stage IIA/B</td></tr>
<tr><td colspan="2">Radiotherapy (standard)</td><td colspan="2">Radiotherapy to paraaortic and ipsilateral iliac field and aditional boost to the enlarged lymph nodes if necessary</td></tr>
<tr><td colspan="2">Chemotherapy
• 3 cycles of BEP or 4 cycles of EP</td><td colspan="2">Chemotherapy
• 3 cycles of BEP or 4 cycles of EP</td></tr>
<tr><td colspan="4">Clinical stage IIC/III</td></tr>
<tr><td>Good prognosis</td><td>Chemotherapy
• 3 cycles of BEP or 4 cycles of EP</td><td>Good prognosis</td><td>Chemotherapy
• 3 cycles of BEP or 4 cycles of EP</td></tr>
<tr><td>Intermediate prognosis</td><td>Chemotherapy
• 4 cycles of BEP or VIP</td><td>Intermediate prognosis</td><td>Chemotherapy
• 4 cycles of BEP</td></tr>
<tr><td colspan="2"></td><td>Poor prognosis</td><td>Chemotherapy
• 4 cycles of BEP</td></tr>
<tr><td colspan="4">AUC: Area under the curve; BEP: Cisplatin, etoposide, bleomycin; EP: Etoposide, cisplatin; RPLND: Retroperitoneal lymph node dissection; VIP: Etoposide, cisplatin, ifosfamide.</td></tr>
</table>

2. EPIGENETICS

The term “Epigenetics” was first described by C.H Waddington in 1942 as “causal interactions between genes and their products, which bring the phenotype into being”. Currently, Epigenetics is defined as the study of phenomena in which heritable and reversible changes in gene expression/function occur without altering the DNA sequences [82, 83].

2.1. Epigenetic Mechanisms

Epigenetic mechanisms include four major groups of modifications: DNA methylation, histone post-translational modifications, histone variants and non-coding RNAs (ncRNAs) (Figure 3) [84]. These mechanisms constitute the epigenome machinery and modulate the chromatin framework (euchromatin – activated transcription or heterochromatin – silenced transcription) which, in turn, culminate in gene expression or repression [84, 85]. Epigenetic mechanisms are essential for normal development as well as for maintenance of specific gene expression profiles in mammals [86]. Nevertheless, dysregulation of these mechanisms are commonly seen in several diseases including cancer, contributing to its development and progression through inactivation of tumor suppressor genes (TSG) or activation of oncogenes [87]. Because of their reversible characteristics, there has been greater interest in studying the epigenetic mechanisms underlying tumorigenesis as well as the search for new therapeutic approaches for cancer treatment [86].

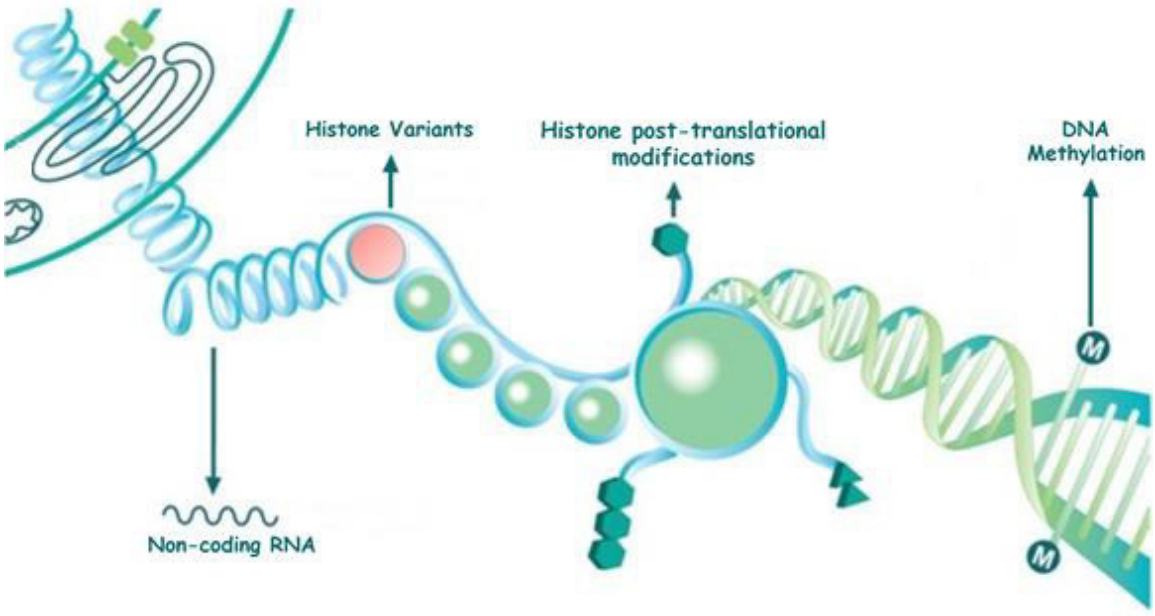


Figure 3. Representation of four different epigenetic mechanisms. M – 5-methylcytosine; the symbols hexagon and triangle represent different translational modifications in N-terminal tails of histones proteins. (CBEG, IPO Porto).

2.1.1. Non-Coding RNAs

Most of the genome is composed by non-protein-coding genes, whose transcription gives rise to ncRNAs consisting in RNA sequences that do not translate into proteins. In fact, only 1.5% of the genome codes into proteins [88, 89]. These ncRNAs are involved in gene expression regulation, messenger RNA (mRNA) degradation, splicing, transport and translation, having an important role in normal development and disease progression [90]. These ncRNAs are generally divided, according to their size, into small non-coding RNAs (sncRNAs) (<200 bp), mid-size non-coding RNAs and long non-coding RNAs (lncRNAs) (>200 bp). sncRNAs include microRNAs (miRNAs) and PIWI-interacting RNAs, mid-size non-coding RNAs include small nucleolar RNAs and lncRNAs comprise large intergenic ncRNAs and transcribed ultraconserved regions [88]. The most well studied ncRNAs are miRNAs [88]. About 60% of human genes are regulated by miRNAs, which bind to complementary sequences at 3'-untranslated region (UTR) of target mRNAs and negatively regulate the mRNA translation into proteins through mRNA degradation or inhibition of translation initiation [88, 91, 92]. A single miRNA may regulate dozens of target mRNAs at the same time and each transcript might be regulated by various miRNAs. This post-transcriptional regulation mediates multiple processes including cell cycle, metabolism, proliferation, differentiation, apoptosis and development. Importantly, miRNAs expression dysregulation has been implicated in various cancers [91, 92]. Thus, miRNAs can act as oncogenes when are overexpressed, or as TSG when downregulated, having also been described as markers for diagnosis, prognosis and as modulators of cancer cells sensitivity to chemotherapeutic agents [93-95].

2.1.2. Histone Post-Translation Modifications and Variants

Chromatin is a DNA packaged state into cells. The basic unit of chromatin is the nucleosome composed by an octamer of four proteins (one pair of each H2A, H2B, H3 and H4), which is wrapped by 147 base pairs of DNA [96]. Histones represent not only the DNA-packaging proteins, but are also recognized as important regulators of chromatin dynamic [97]. Histones undergo a series of post-translational modifications such as acetylation, methylation and phosphorylation, influencing the chromatin conformation [98]. The set of different changes that may occur in histone is called "histone code", which is important in various processes such as gene expression and DNA replication and repair [97, 98]. Furthermore, excepting H4, the other histones that constitute the nucleosome (H2A, H2B and H3) have several histone variants, contributing to more diversification of chromatin conformation, important for cellular regulation during development. These variants,

contrarily to their canonical counterparts are expressed independently of DNA replication [99, 100].

2.1.3. DNA Methylation

DNA methylation is the most widely studied epigenetic mechanism [86, 87], occurring almost exclusively in cytosine residues that precedes guanines in structures named CpG dinucleotides (Cytosine-phosphate-Guanine) in mammals [101, 102]. This process is catalyzed by a set of enzymes called DNA methyltransferases (DNMTs), using S-adenosyl-methionine (SAM) as a methyl group (CH_3) donor and consists in adding the methyl group to the 5'carbon of cytosine, originating 5-methylcytosine (5mC) (Figure 4) [87, 103]. DNMTs family includes 5 members: DNMT1, DNMT2 and DNMT3 (DNMT3A, DNMT3B and DNMT3L), wherein only DNMT1, DNMT3A and DNMT3B possess methyltransferase activity. DNMT2 is involved in methylation of RNA [104, 105]. DNMT1 is a maintenance methyltransferase responsible for propagation of original methylation patterns during somatic cell divisions, initially established by *de novo* methyltransferases. DNMT1 preferentially acts on hemi-methylated DNA during replication. DNMT3A and DNMT3B are *de novo* methyltransferases responsible for establishment of DNA methylation patterns during embryogenesis, that act independently of replication on unmethylated and hemi-methylated DNA [85, 86, 106]. Unlike DNMT3A and DNMT3B, DNMT3L is catalytically inactive since it lacks the catalytic domain that binds to SAM, but acts as a regulator factor that interact with catalytic domains of DNMT3A and DNMT3B stimulating their function as *de novo* DNMTs [107].

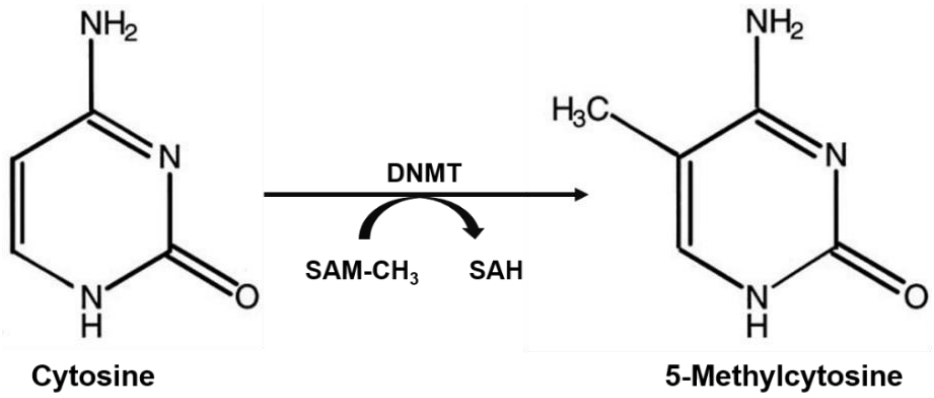


Figure 4. DNA methylation reaction: conversion of a cytosine to 5-methylcytosine catalyzed by DNA methyltransferases (DNMTs). These enzymes promote the addition of a methyl group (CH_3) from S-adenosyl-methionine (SAM) to 5'carbon of cytosine. SAH: S-adenosyl-homocysteine.

In human genome about 75% of all CpG dinucleotides are methylated [104]. CpG sites are not randomly distributed in genome, since these are especially clustered in short CpG-dense stretches named CpG islands and in regions of large non-coding repetitive

sequences such as centromeres and retrotransposons [86]. CpG island is defined as a region with more than 200 base pairs, G+C content of at least 50% and observed-to-expected CpG frequencies ratio of, at least, 0.6 [105]. CpG islands are preferentially located in transcriptional start sites of human genes, occupying about 50-60% of human genes promoters, which show the important role of these regions in transcriptional regulation [106, 108].

In normal mammalian cells, the CpG dinucleotides within CpG islands are usually unmethylated, maintaining an open state of chromatin, allowing gene transcription (Figure 5). Nevertheless, some of these regions become methylated in a tissue-specific manner during development, leading to gene silencing such as the epigenetic silencing of chromosome X and in imprinted genes [86, 104, 108]. Furthermore, DNA methylation may also occur in CpG islands shores closely located to CpG islands, which are also unmethylated in normal human cells [105]. In contrast, the CpG dinucleotides found in non-coding repetitive elements and gene bodies are often heavily methylated preventing chromosomal instability (through silencing non-coding DNA and transposable DNA elements) and avoiding transcription initiation in incorrect sites, respectively (Figure 5) [105].

DNA methylation is involved in several cellular processes and its dysregulation has been associated with the development of cancer, being involved in initiation, progression, invasion and metastases formation. There are two main altered methylation patterns in malignancies, namely global DNA hypomethylation and local hypermethylation of specific CpG islands, mainly in TSG promoters (Figure 5) [87, 105]. Indeed, 20-60% of tumor cells are characterized by global loss of methylation when compared to normal cells [105]. DNA hypomethylation was one of the first epigenetic alterations associated with cancer and occurs particularly in repetitive sequences, retrotransposons, coding regions and introns that are usually methylated [102]. The consequences associated with loss of methylation include chromosomal instability, reactivation of genes implicated in cancer, oncogenes activation and loss of imprinting [86, 102]. TSG are involved in various biological functions including cell cycle, DNA repair, cellular interactions, apoptosis and angiogenesis. Promoter hypermethylation in these genes result in their inadequate inactivation and constitutes the major event of cancer's initiation and progression [102].

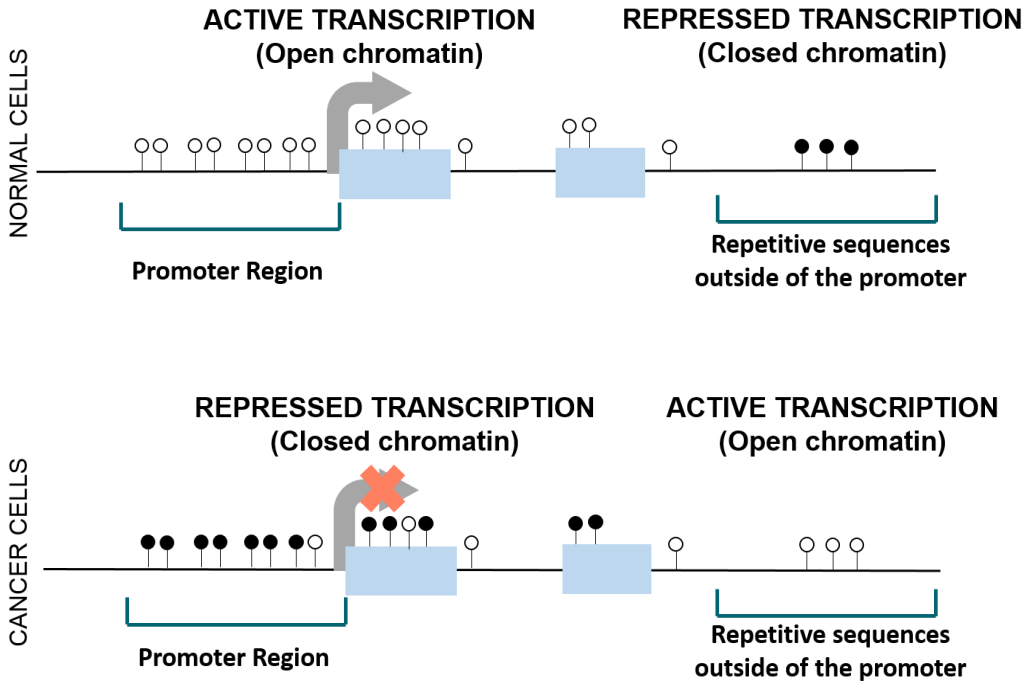


Figure 5. DNA methylation in normal and cancer cells. In normal cells, CpG islands localized in the genes' promoter are unmethylated (cytosine, *open lollipops*), maintaining an open conformation of chromatin (euchromatin state) and, consequently, allowing gene transcription (*large gray arrow*). In contrast, the CpG dinucleotides localized outside the promoter are heavily methylated (cytosine, *black lollipops*), characterized by a closed conformation of chromatin (heterochromatin) and therefore, transcription inactivation. In cancer cells the reverse pattern of normal cells is observed: CpG islands of genes' promoter become abnormally methylated (cytosine, *black lollipops*), leading to transcriptional silencing (*large gray arrow with red X*), while the regions outside of promoter become unmethylated (cytosine, *open lollipops*), allowing gene transcription. Blue box represent exons.

DNA methylation is associated with a repressive state of chromatin, resulting in transcription inhibition and, consequently, gene expression repression. Gene silencing can occur either by directly blocking the binding of transcription factors to their specific DNA target sites and/or indirectly through recruitment of methyl-CpG-binding (MBD) proteins (MBD1, MBD2, MBD3, MBD4, methylcytosine-binding protein 2 (MeCP2) and kaiso), which recognize and bind to methylated CpG sites [105, 108]. These proteins may recruit chromatin remodeling co-repressor complexes such as histone deacetylases and histone methyltransferases which, in turn, may induce chromatin remodeling, important in gene silencing [109].

2.2. Epigenetic Alterations in Testicular Germ Cell Tumors

The foremost epigenetic alterations described for TGCT have already been compiled in a review of the literature, published in *Epigenomics* 2017 [Costa AL *et al.* "The Epigenetics of Testicular Germ Cell Tumors: Looking for Novel Disease Biomarkers" (Annex I)] [110].

Particularly, in this study we focus in DNA methylation dysregulation, specifically in *CRIPTO*, *HOXA9*, *MGMT*, *RASSF1A* and *SCGB3A1* genes. Five genes localization in genome and biological function were stated in Table 6.

Table 6. Five genes localization in genome and biological function.

Gene	Locus	Function	Reference
CRIPTO	3p21.31	Essential during embryogenesis; involved in stem cell self-renewal and pluripotency in human embryonic stem cells	[111]
HOXA9	7p15.2	Homeobox family genes are involved in regulation of morphogenesis, cell differentiation and determination of cell identity	[112]
MGMT	10q26.3	Involved in DNA repair, whose protein protects cells from promutagenic O6-methylguanine alkylation by removing of DNA adducts formed by alkylating agents	[113]
RASSF1A	3p21.31	TSG involved in multiple biological pathways such as apoptosis, cell motility and invasion and cell cycle regulation	[114]
SCGB3A1	5q35.3	TSG involved in apoptosis, cell cycle control and suppression of cell migration and invasion	[115]

II. AIMS OF STUDY

TGCT are a heterogeneous group of tumors, although most of them share the same cytogenetic and genetic background, as stated above. Furthermore, the classical serum tumor markers have low sensitivity and specificity for TGCT detection, with only 60% of TGCT patients exhibiting elevated levels of these markers at diagnosis. Nonetheless, epigenetic alterations that underlie the diversity of TGCT phenotypes and behavior, might constitute potential biomarkers for TGCT and provide additional information to standard diagnosis tools (clinical, serological, imagiological and pathological parameters). Importantly, *CRIPTO*, *HOXA9*, *MGMT*, *RASSF1A* and *SCGB3A1* promoter methylation has been reported for these tumors, as described in the literature review included in Annex I.

Therefore, the main goal of this master dissertation is to evaluate the methylation pattern of five genes promoters among TGCT subtypes and assess their potential as methylation-based biomarkers for subtype discrimination, as well as identify biological differences among TGCT subtypes, thus explaining its heterogeneity.

Specifically, the major aims are:

- Validate a five-genes (*CRIPTO*, *HOXA9*, *MGMT*, *RASSF1A* and *SCGB3A1*) promoter methylation panel in a retrospective cohort of TGCT and normal tissues to assess its value as biomarker of neoplasia;
- Assess whether specific genes methylation profiles are associated with different histological TGCT subtypes and how they do correlate with standard clinicopathological parameters;
- Evaluate the prognostic value of the promoter methylation levels of the five genes in TGCT.

III. MATERIALS AND METHODS

1. PATIENTS AND TISSUE SAMPLE COLLECTION

Formalin-fixed paraffin-embedded (FFPE) tissue samples from 161 cases of primary TGCT [146 related to GCNIS - 76 pure SE and 70 NST (10 pure EmbrCa, 4 pure postpubertal-type TE, and 56 mixed germ cell tumors) – and 15 unrelated to GCNIS - 4 prepubertal-type YST and 11 prepubertal-type TE], diagnosed between 2005 and 2016 and archived at the Department of Pathology at Portuguese Oncology Institute of Porto, Portugal were selected for this study. All the components of mixed germ cell tumors were identified in representative histological slides (see below), selectively collected and individually considered for the purposes of comparisons among TGCT subtypes. Thus, malignant TGCT tissue samples included 238 samples of TGCT related to GCNIS – 76 pure SE and 162 NST (10 pure EmbrCa and 4 pure postpubertal-type TE, as well as 19 SE, 42 EmbrCa, 36 postpubertal-type YST, 14 CH and 37 postpubertal-type TE as components of mixed germ cell tumors).

All tumor samples were collected from orchiectomy specimens of TGCT patients, prior to any systemic treatment, and were routinely fixed and paraffin-embedded for routine histopathologic diagnosis and staging. Representative blocks were selected by a pathologist in slides previously stained with hematoxylin and eosin (H&E), and relevant areas were identified for subsequent DNA extraction. Additionally, 16 non-malignant testicular FFPE tissue samples were collected from patients submitted to orchiectomy due to testicular inflammatory disease or benign pathology and were used as controls. Relevant clinicopathological data of TGCT patients and controls were collected from clinical charts. This study was approved by institutional ethics review board (Comissão de Ética para a Saúde – CES-IPO-12-017) of Portuguese Oncology Institute of Porto, Portugal.

2. DNA EXTRACTION FROM FORMALIN-FIXED PARAFFIN-EMBEDDED TISSUES (FFPE)

DNA was extracted from FFPE tissue samples using FFPE RNA/DNA Purification Plus Kit (Norgen, Canada, USA), according to manufacturers' recommendations with some adjustments (detailed protocol in Annex II). Briefly, sections of 10 µm were cut from FFPE tissue block, previously selected by a pathologist as mentioned before, using a microtome. Tumor areas of interest were macrodissected from these 10 µm sections using a disposable sterile scalpel blade and transferred to 1.5 mL safe-lock tubes. The first step consisted on deparaffinization of FFPE samples through a series of xylene (VWR Chemicals, France) and absolute ethanol (Merck, Germany) washes. Next, the FFPE samples were digested by a specific time of incubation with Proteinase K (20 mg/mL) (Nzytech, Portugal) and Digestion Buffer A. Then, the lysate containing RNA was collected for RNA purification

following a specific protocol, while the pellet containing DNA was further digested for DNA extraction.

Focusing on DNA extraction, Buffer RL and absolute ethanol (Merck, Germany) were added to DNA-containing lysate and the solution was loaded onto a DNA Purification Micro Column. The last step consists in a series of washes with the provided Wash Solution A. Lastly, the purified DNA was eluted in 20 µL Elution Buffer F provided.

After elution, for each sample, the DNA concentration and respective purification ratios were assessed by NanoDrop Lite Spectrophotometer (NanoDrop Technologies, USA) and stored at -20°C.

3. SODIUM BISULFITE TREATMENT OF DNA

Extracted DNA was treated with sodium bisulfite, which consists in a chemical reaction where the unmethylated cytosines are converted to uracils, while methylated cytosines remains cytosines (Figure 6). Thus, sodium bisulfite treatment of DNA allows discrimination between methylated and unmethylated DNA sequences, which can be further analyzed by polymerase chain reaction (PCR)-based methods.

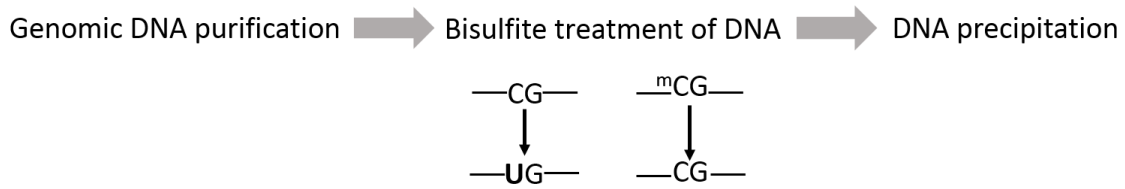


Figure 6. Sodium bisulfite treatment of genomic DNA. In this reaction, the unmethylated cytosines are converted to uracils, while methylated cytosines remain unchanged.

Sodium bisulfite treatment of extracted DNA was performed using EZ DNA Methylation-Gold™ Kit (Zymo Research, CA, USA), according to manufacturers' recommendations (detailed protocol in Annex III). The quantity of DNA used was variable according to DNA concentration obtained from each sample. Thus, the quantity of bisulfite-modified genomic DNA varied between 250-1200 nanograms and consequently, it was eluted in 12.5-60 µL of sterile bidistilled water (B.Braun, Melsungen, Germany) and stored at -80°C.

CpGenome™ Universal Methylated DNA (Merck Millipore, Germany) was also modified using the same process for further use in quantitative methylation-specific PCR (qMSP) as positive control. In the last step, it was eluted in a total volume of 20 µL of sterile bidistilled water (B.Braun, Melsungen, Germany).

4. QUANTITATIVE METHYLATION-SPECIFIC POLYMERASE CHAIN REACTION (QMSP)

To evaluate the degree of each gene methylation, qMSP method was performed. This method allows a specific, sensitive, cost-effective and high-throughput analysis [116].

In qMSP, specific primers were used to amplify methylated dinucleotides CpG. For *β-Actin*, *MGMT* and *SCGB3A1* genes, primers were selected from the literature [117-119] (Table 7). On the other hand, for *CRIPTO*, *HOXA9* and *RASSF1A* genes, sequence-specific primers were designed using *in silico* PCR tools and synthesized by Sigma Aldrich. The dye used to detect DNA amplification was SyBr Green, which emits a fluorescent signal that is proportional to the quantity of amplified PCR product. Although, this dye is non-specific because it binds to double strand DNA (dsDNA), thus, a melting curve was generated in order to assess the presence of non-specific PCR products and primer-dimer [120].

Bisulfite-modified genomic DNA was used as template to evaluate the methylation status of *CRIPTO*, *HOXA9*, *MGMT*, *RASSF1A* and *SCGB3A1* genes and reactions were carried out in 384-well plates using LightCycler 480 (Roche, Germany). In brief, 1 µL of modified DNA, 5 µL of NZYSpeedy qPCR Green Master Mix (2x), ROX (Nzytech, Portugal), a variable volume of primers depending of gene analyzed (Table 7) and sterile bidistilled water (B.Braun, Melsungen, Germany), were added in a total volume of 10 µL in each well. The thermocycler conditions consisted in an initial step of polymerase activation at 95°C for 2 minutes followed by 40 cycles at 95°C for 5 seconds and 20 seconds at a variable temperature, according to the different primers used (Table 7), for denaturation and annealing/extension, respectively. Each sample was carried out in triplicate and, in each plate two negative controls and five serial dilutions (in duplicate) of a previously modified CpGenome™ Universal Methylated DNA (Merck Millipore, Germany) were also included. Based on these dilutions, a standard curve was generated allowing relative quantification and ascertain the PCR efficiency.

In addition, *β-Actin* gene was used as internal reference gene for the normalization of the assay. The relative methylation level was determined by comparison the mean concentration of each target gene for each sample and the mean concentration of *β-Actin* gene for each sample according to the next formula:

$$\textbf{Methylation level} = \frac{\textbf{Target gene}}{\boldsymbol{\beta - Actin}} \times \mathbf{1000}$$

Table 7. Primers sequences and qMSP conditions for each of the analyzed genes.

Gene	Primer sequence 5'-3'	Amplicon (bp)	Annealing Temperature (°C)	Concentration per reaction (F+R)	Reference
β-Actin	**F:**TGGTGATGGAGGAGGTTTAGTAAGT **R:**AACCAATAAAACCTACTCCTCCCTTAA	133	60	400 nM	[117]
CRIPTO	**F:**GAGGCGATTTCGGTTTATAGAC **R:**CTAACCCCACGACACCGA	134	60	200 nM	n.a
HOXA9	**F:**TATTTAGTCGGTATTCGC **R:**ACCTCGAACGCTTCCAT	150	60	300 nM	n.a
MGMT	**F:**TTTCGACGTTCGTAGGTTTTCGC **R:**GCACTCTTCCGAAAACGAAACG	81	62	400 nM	[118]
RASSF1A	**F:**AGCGAAGTACGGGTTTAATC **R:**ACACGCTCCAACCGAATA	111	60	300 nM	n.a
SCGB3A1	**F:**GGTACGGGTTTTTTACGGTTCGTC **R:**AACTTCTTATACCCGATCCTCG	135	64	400 nM	[119]

n.a: Not applicable.

5. STATISTICAL ANALYSIS

Differences in the five genes promoter methylation levels for all combinations of TGCT subtypes and comparisons with controls were assessed by non-parametric Kruskal-Wallis test for multiple comparisons. Pairwise comparisons were carried out using non-parametric Mann-Whitney U test with Bonferroni's correction (*i.e.*, dividing the standard p-value by the number of the groups included in the multiple comparison analysis). Non-parametric Mann-Whitney U Test was also used for seeking associations between promoter methylation

levels and clinicopathological parameters. In patients diagnosed with mixed TGCT, the highest methylation value of mixed components was considered for comparing genes promoter methylation levels and clinicopathological features.

The biomarker performance of each gene and of the different combinations of genes was assessed through receiver operating characteristics (ROC) curve construction. For each gene, a ROC curve was constructed plotting sensitivity (true positive) against 1-specificity (false positive) whereas for genes-panel, a logistic regression model was used. For each gene, a cut-off was established by ROC curve [sensitivity + (1-specificity)], to maximize the sensitivity and specificity. In addition, area under the curve (AUC), sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and accuracy were ascertained for each gene individually and for the different genes-panels (Table 8). For this analysis, a given genes-panel methylation test was considered positive when at least one of the genes tested was positive individually. Moreover, methylation frequencies were determined for each TGCT subtype using the highest (for hypermethylation) and lowest (for hypomethylation) values of controls as methylation cut-offs for each gene.

Five genes methylation levels were correlated with patients' age using a Spearman non-parametric correlation test.

Globally, results were considered statistically significant when p-value<0.05. Statistical analysis was performed using SPSS Statistics, version 24 (IBM-SPSS, IL, USA) and GraphPad Prism 6 (GraphPad Software, CA, USA).

Table 8. Biomarkers performance calculation for each gene.

	TGCT samples	Controls
Positive (> cut-off)	A	B
Negative (< cut-off)	C	D
Total (n)	E	F

Genes Efficacy	Formula (%)
Sensitivity	$\frac{A}{E}$ X 100
Specificity	$\frac{D}{F}$ X 100
PPV	$\frac{A}{A+B}$ X 100
NPV	$\frac{D}{C+D}$ X 100
Accuracy	$\frac{A+D}{E+F}$ X 100

PPV: Positive predictive value; NPV: Negative predictive value.

IV. RESULTS

1. CLINICAL AND PATHOLOGICAL DATA

All clinical samples from orchiectomy specimens included in this study were treatment naive. The median age of patients included was 30 years (range 13 to 52) for GCNIS-related TGCT and 16 years (range 1 to 35) for GCNIS-unrelated TGCT. The majority of GCNIS-related cases (52.1%) presented SE (76/146), whereas 47.9% presented NST (70/146), being 14.3% of EmbrCa (10/70), 5.7% of postpubertal-type TE (4/70) and 80.0% of mixed tumors (56/70). Concerning GCNIS-unrelated TGCT, 73.3% of cases presented prepubertal-type TE (11/15) followed by prepubertal-type YST (4/15). Most of patients (63.7%) with GCNIS-related TGCT presented clinical stage I (93/146), whereas only 11.0% (16/146) were stage III. According to the International Germ Cell Cancer Collaborative Group (IGCCCG) [78], 26.8% (37/138) of the cases were classified at good prognosis, 2.2% (3/138) at intermediate prognosis and 3.6% (5/138) at poor prognosis.

Before surgery, increased levels of AFP, β-hCG and LDH were observed in 35.0% (49/140), 51.8% (72/139) and 42.9% (48/112) of GCNIS-related cases, respectively (Supplementary Table – Annex IV).

The most relevant clinical and pathological characteristics of our cohort are described in Table 9.

Table 9. Clinical and pathological characteristics of TGCT patients and controls included in this study.

CLINICOPATHOLOGIC CHARACTERISTICS	TGCT PATIENTS		CONTROLS
	GCNIS-related cases	GCNIS-unrelated cases	
Patients, n	146	15	16
Median age, years (range)	30 (13-52)	16 (1-35)	52 (16-86)
Histologic subtype			
Seminoma	76/146	n.a	n.a
Embryonal carcinoma	10/146	n.a	n.a
Postpubertal-type yolk-sac	0	n.a	n.a
Choriocarcinoma	0	n.a	n.a
Postpubertal-type teratoma	4/146	n.a	n.a
Mixed germ cell tumors	56/146	n.a	n.a
Prepubertal-type teratoma	n.a	11/15	n.a
Prepubertal-type yolk-sac	n.a	4/15	n.a
Group staging			
I	93/146	n.a	n.a
II	37/146	n.a	n.a
III	16/146	n.a	n.a
Prognostic-based staging (according IGCCCG classification), for metastatic disease*			
Not applicable (Non-metastatic)	93/138	n.a	n.a
Good	37/138	n.a	n.a
Intermediate	3/138	n.a	n.a
Poor	5/138	n.a	n.a

TGCT: Testicular germ cell tumors; GCNIS: Germ cell neoplasia *in situ*; IGCCCG: International Germ Cell Cancer Collaborative Group;
*The information about prognostic-based staging were not available in some of TGCT cases;
n.a: not applicable.

2. ASSESSMENT OF THE DIAGNOSTIC PERFORMANCE OF GENE PROMOTER METHYLATION

CRIPTO, *HOXA9*, *MGMT* and *SCGB3A1* methylation levels were significantly higher in GCNIS-related TGCT compared to controls (p=0.0058, p=0.0092, p=0.0268 and p=0.0156, respectively) (Figure 7). However, the diagnostic performance of these markers was modest, with AUC values of 0.7033 for *CRIPTO*, 0.6922 for *HOXA9*, 0.6282 for *MGMT* and 0.6685 for *SCGB3A1*.

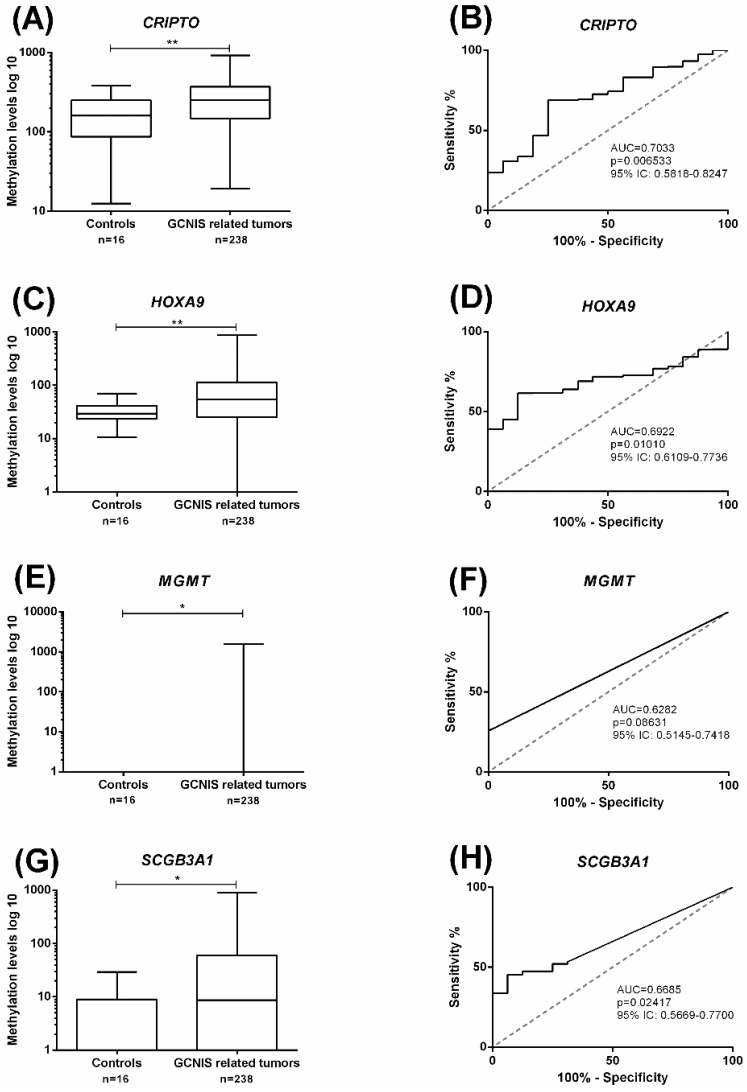


Figure 7. Distribution and ROC curve analyses of four genes promoter methylation levels in controls and TGCT related to GCNIS. Left panel: Boxplots of (A) *CRIPTO*, (C) *HOXA9*, (E) *MGMT* and (G) *SCGB3A1* promoter methylation levels in controls and TGCT related to GCNIS. (Mann-Whitney U Test, *p<0.05, **p<0.01). Right panel: ROC curve for (B) *CRIPTO*, (D) *HOXA9*, (F) *MGMT* and (H) *SCGB3A1* across controls and TGCT related to GCNIS. (Dark line: ROC curve for *CRIPTO*, *HOXA9*, *MGMT* and *SCGB3A1* genes; light gray dashed line: reference line).

Conversely, no significant differences between controls and GCNIS-related TGCT were found for *RASSF1A* methylation levels.

Notably, comparisons between pure SE samples *vs.* controls and between NST samples *vs.* controls, disclosed that *CRIPTO*, *HOXA9*, *MGMT* and *SCGB3A1* methylation levels were significantly higher in NST ($p=0.0002$, $p<0.0001$, $p=0.0044$ and $p<0.0001$, respectively), whereas *RASSF1A* promoter methylation levels were significantly higher in controls compared with pure SE tissues ($p<0.0001$) (Table 10 and Figure 8).

Table 10. *CRIPTO*, *HOXA9*, *MGMT*, *RASSF1A* and *SCGB3A1* promoter methylation levels in controls and GCNIS-related TGCT.

Genes	Controls Median (IQR)	SE Median (IQR)	p-value Controls *vs.* SE	NST Median (IQR)	p-value Controls *vs.* NST
CRIPTO	159.26 (86.45-249.42)	172.43 (104.93-257.14)	0.5050	315.39 (197.91-421.81)	0.0002
HOXA9	29.49 (23.42-41.34)	22.29 (10.18-42.15)	0.1346	83.94 (45.45-144.71)	<0.0001
MGMT	0 (0-0)	0 (0-0)	>0.9999	0 (0-48.82)	0.0044
RASSF1A	14.21 (7.00-26.41)	0 (0-0)	<0.0001	28.95 (1.20-102.95)	0.1101
SCGB3A1	0 (0-8.93)	0 (0-0)	0.2802	25.58 (0-84.43)	<0.0001

SE: Seminomas; NST: Nonseminomatous tumors; IQR: Interquartile range.

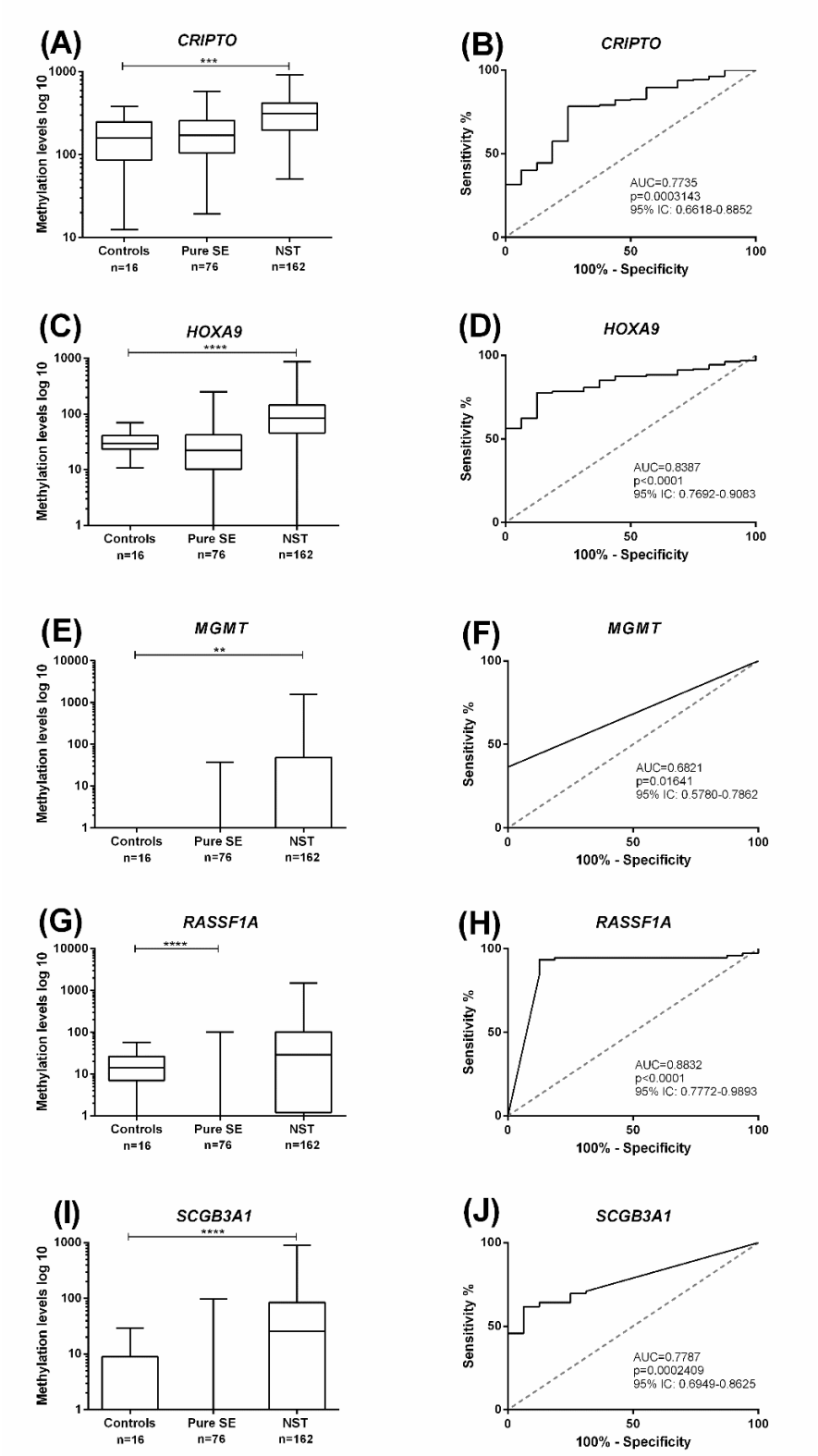


Figure 8. Distribution and ROC curve analyses of five genes promoter methylation levels in controls and pure SE or NST samples. Left panel: Boxplots of (A) *CRIPTO*, (C) *HOXA9*, (E) *MGMT*, (G) *RASSF1A* and (I) *SCGB3A1* promoter methylation levels in controls, pure SE and NST samples. (Mann-Whitney U Test, **p<0.01, ***p<0.001, ****p<0.0001). Right panel: ROC curve for (B) *CRIPTO*, (D) *HOXA9*, (F) *MGMT* and (J) *SCGB3A1* across controls and NST samples; ROC curve for (H) *RASSF1A* across controls and pure SE samples. (Dark line: ROC curve for five genes; light gray dashed line: reference line).

CRIPTO, *HOXA9* and *SCGB3A1* promoter methylation levels displayed AUC values above 0.75 for discriminating NST from controls (Table 11). Remarkably, *HOXA9* showed the highest values for all standard validity estimates, except for NPV (28%) (Table 11).

Moreover, the combination of *CRIPTO*/*HOXA9*/*SCGB3A1* panel improved the diagnostic coverage of NST, with high sensitivity (97.5%), PPV (96.9%) and accuracy (94.9%), despite modest specificity (69%) (Table 12 and Figure 9).

Table 11. Performance of gene promoter methylation for NST identification.

Genes	AUC	Sensitivity %	Specificity %	PPV %	NPV %	Accuracy %
CRIPTO	0.7735	78.4	75	96.9	25.5	78.1
HOXA9	**0.8387**	**77.8**	**87.5**	**98.4**	**28**	**78.7**
SCGB3A1	0.7787	61.7	93.8	99	19.5	64.6
AUC: Area under the curve; PPV: Positive predictive value; NPV: Negative predictive value.						

Table 12. Performance of genes-panel methylation for NST identification.

Genes	AUC	Sensitivity %	Specificity %	PPV %	NPV %	Accuracy %
CRIPTO/HOXA9	0.889	95.1	75	97.5	60	93.3
CRIPTO/SCGB3A1	0.872	92	68.8	96.8	45.8	89.9
HOXA9/SCGB3A1	0.894	87.7	81.3	97.9	39.4	87.1
CRIPTO/HOXA9 /SCGB3A1	**0.915**	**97.5**	**68.8**	**96.9**	**73.3**	**94.9**
AUC: Area under the curve; PPV: Positive predictive value; NPV: Negative predictive value.						

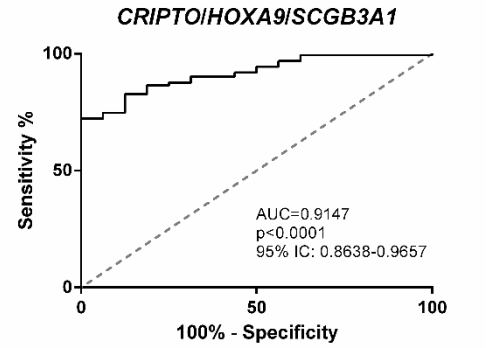


Figure 9. ROC curve for *CRIPTO*/*HOXA9*/*SCGB3A1* panel across controls and NST samples. (Dark line: ROC curve for *CRIPTO*/*HOXA9*/*SCGB3A1* genes-panel; light gray dashed line: reference line).

Concerning discrimination from pure SE, *RASSF1A* displayed AUC value of 0.8832 (Figure 8H), corresponding to 87.5% sensitivity, 93.4% specificity, 97.3% NPV and 92.4% accuracy, although PPV was only 74%.

3. PROMOTER METHYLATION PROFILE IN TGCT SUBTYPES

3.1. Pure Seminomas *versus* Nonseminomatous Tumors

Regarding the two major subtypes of GCNIS-related TGCT, *CRIPTO*, *HOXA9*, *MGMT*, *RASSF1A* and *SCGB3A1* promoter methylation levels were significantly higher in NST compared to pure SE ($p<0.0001$ for all genes, Table 13 and Figure 10).

Table 13. *CRIPTO*, *HOXA9*, *MGMT*, *RASSF1A* and *SCGB3A1* promoter methylation levels in GCNIS-related TGCT.

TGCT SUBTYPES			GENES				
			CRIPTO	*HOXA9*	*MGMT*	*RASSF1A*	*SCGB3A1*
Pure SE Median (IQR)			172.43 (104.93-257.14)	22.29 (10.18-42.15)	0 (0-0)	0 (0-0)	0 (0-0)
NST Median (IQR)	Mixed SE		291.68 (216.86-363.44)	27.50 (13.55-38.61)	0 (0-0)	0 (0-13.50)	0 (0-11.45)
	EmbrCa	Pure	189.07 (125.61-290.25)	87.29 (59.48-114.04)	0 (0-0)	2.25 (0.59-19.40)	21.68 (4.91-28.33)
		Mixed	279.07 (162.54-415.15)	90.61 (51.00-131.93)	0 (0-27.83)	20.03 (1.20-57.78)	17.72 (0-50.10)
	Mixed YST		276.28 (149.98-417.25)	88.98 (60.46-260.73)	0 (0-47.46)	82.01 (18.18-325.69)	17.72 (0-50.10)
	Mixed CH		337.23 (240.57-468.27)	72.20 (55.44-183.19)	0 (0-13.69)	128.53 (30.05-275.40)	85.10 (45.77-168.39)
	TE	Pure	371.65 (365.81-445.34)	37.81 (10.61-103.50)	0 (0-207.40)	0.19 (0-25.94)	0 (0-12.64)
		Mixed	352.67 (234.27-610.60)	117.46 (60.79-169.11)	15.87 (0-147.70)	59.69 (7.59-115.60)	60.02 (10.88-123.09)

SE: Seminomas; NST: Nonseminomatous tumors; IQR: Interquartile range; EmbrCa: Embryonal carcinoma; YST: Postpubertal-type yolk-sac tumor; CH: Choriocarcinoma; TE: Postpubertal-type teratoma; Mixed: TGCT subtypes as components of mixed germ cell tumors.

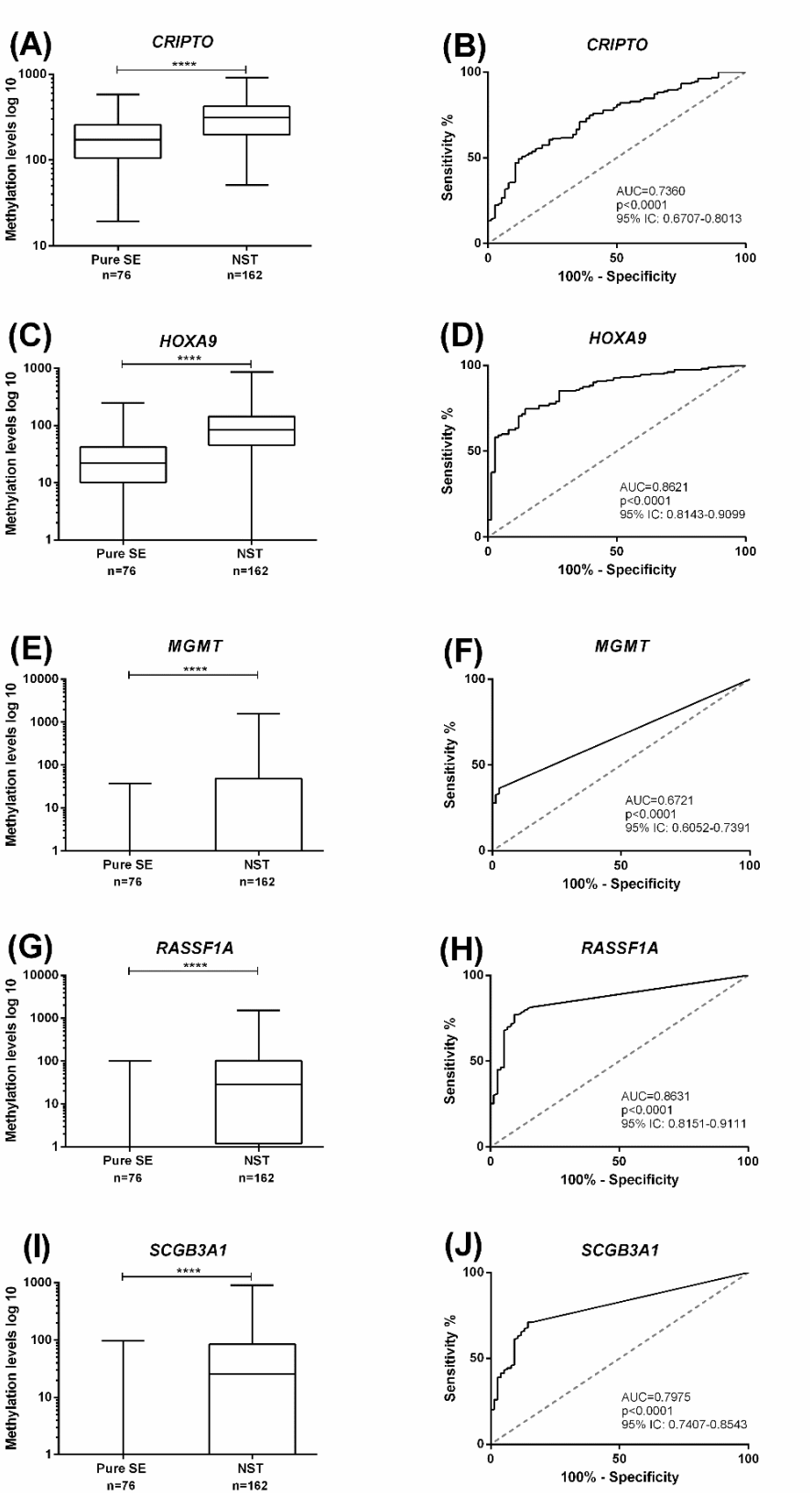


Figure 10. Distribution and ROC curve analyses of five genes promoter methylation levels in pure SE and NST samples. Left panel: Boxplots of (A) *CRIPTO*, (C) *HOXA9*, (E) *MGMT*, (G) *RASSF1A* and (I) *SCGB3A1* promoter methylation levels in pure SE and NST samples. (Mann-Whitney U Test, ****p<0.0001). Right panel: ROC curve for (B) *CRIPTO*, (D) *HOXA9*, (F) *MGMT*, (H) *RASSF1A* and (J) *SCGB3A1* across pure SE and NST samples. (Dark line: ROC curve for five genes; light gray dashed line: reference line).

Importantly, for discriminating pure SE from NST, *RASSF1A* displayed the highest AUC (0.8631), sensitivity, specificity, and accuracy (Table 14). Since *CRIPTO* and *MGMT* promoter methylation displayed AUC values inferior to 0.75, validity estimates were not determined.

Table 14. Discriminatory performance of genes promoter methylation between pure SE and NST samples.

Genes	AUC	Sensitivity %	Specificity %	PPV %	NPV %	Accuracy %
HOXA9	0.8621	74.7	85.5	91.7	61.3	78.2
RASSF1A	**0.8631**	**77.2**	**90.8**	**94.7**	**65.1**	**81.5**
SCGB3A1	0.7975	71	85.5	91.3	58	75.6
AUC: Area under the curve; PPV: Positive predictive value; NPV: Negative predictive value.						

Several combinations of these three-genes were tested to assess the best performance for distinguishing pure SE from NST. Although all combinations displayed an AUC value higher than 0.85 (Table 15), the best overall performance was achieved by *HOXA9*/*RASSF1A* panel with 90.1% sensitivity, 81.6% specificity, 91.3% PPV, 79.5% NPV and 87.4% accuracy (Table 15 and Figure 11).

Table 15. Discriminatory performance of genes-panel methylation between pure SE and NST samples.

Genes	AUC	Sensitivity %	Specificity %	PPV %	NPV %	Accuracy %
HOXA9/RASSF1A	**0.904**	**90.1**	**81.6**	**91.3**	**79.5**	**87.4**
HOXA9/SCGB3A1	0.896	88.3	75	88.3	75	84
RASSF1A/SCGB3A1	0.872	84.6	82.9	91.3	71.6	84
HOXA9/RASSF1A /SCGB3A1	0.911	91.4	73.7	88.1	80	85.7
AUC: Area under the curve; PPV: Positive predictive value; NPV: Negative predictive value.						

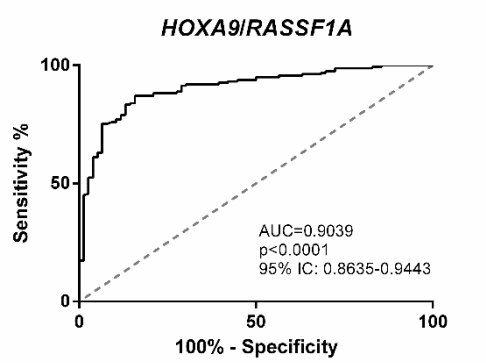


Figure 11. ROC curve for *HOXA9*/*RASSF1A* panel across pure SE and NST samples. (Dark line: ROC curve for *HOXA9*/*RASSF1A* genes-panel; light gray dashed line: reference line).

Concerning aberrant (hyper- or hypo-) promoter methylation frequencies, CH as component of mixed TGCT depicted the highest values for *SCGB3A1* (12/14), *RASSF1A* (10/14) and *HOXA9* (7/14), followed by postpubertal-type TE as component of mixed TGCT with *HOXA9* (27/37), *SCGB3A1* (22/37) and *MGMT* (21/37). Importantly, *RASSF1A* hypomethylation was observed in 64 out of 76 pure SE (Figure 12).

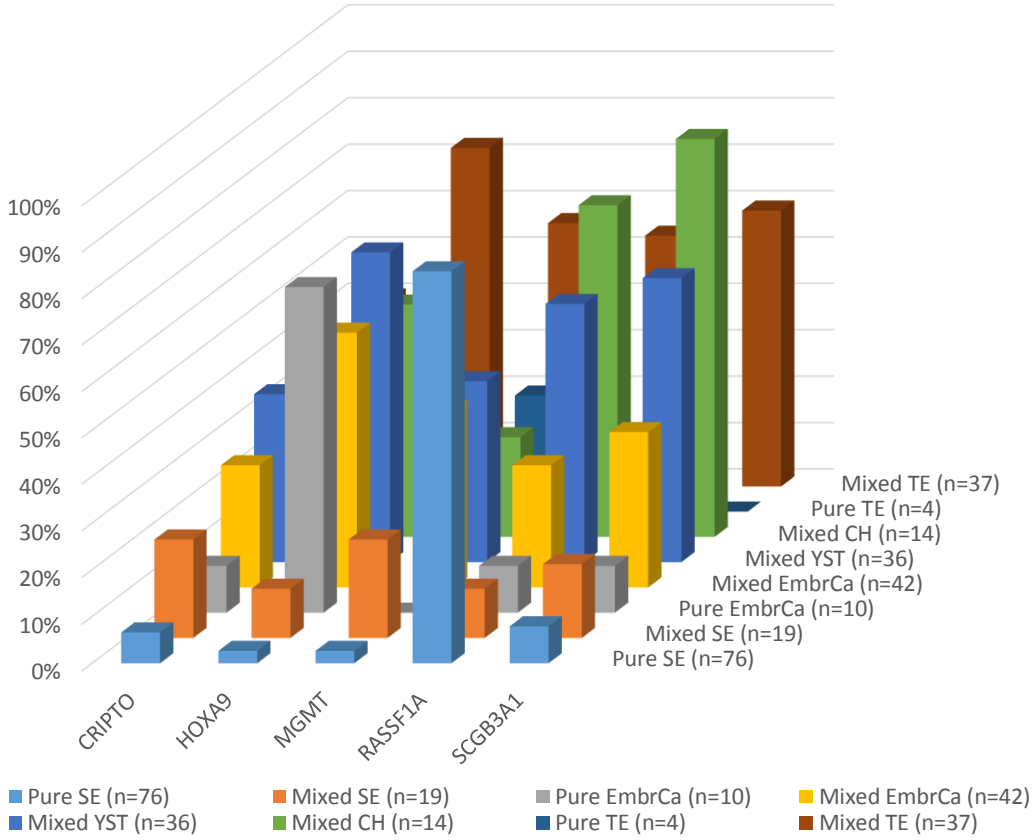


Figure 12. Aberrant promoter methylation frequencies amongst TGCT histological subtypes. (Mixed: TGCT subtypes as components of mixed germ cell tumors; YST: Postpubertal-type YST; TE: Postpubertal-type TE).

3.2. Pure Tumors *versus* Matching Mixed Forms

Interestingly, *CRIPTO*, *HOXA9*, *MGMT*, *RASSF1A* and *SCGB3A1* methylation levels were significantly different in pure GCNIS-related TGCT comparing with their matching mixed forms. Indeed, components of mixed TGCT displayed the highest levels for all studied genes. No patients with pure postpubertal-type YST and CH were available to perform the same type of analysis.

3.2.1. Pure Seminoma *versus* Seminoma in Mixed Germ Cell Tumors

CRIPTO, *MGMT* and *RASSF1A* methylation levels were significantly higher in SE as component of mixed TGCT compared with pure SE (p=0.0004, p=0.0061, p=0.0220, respectively) (Figure 13).

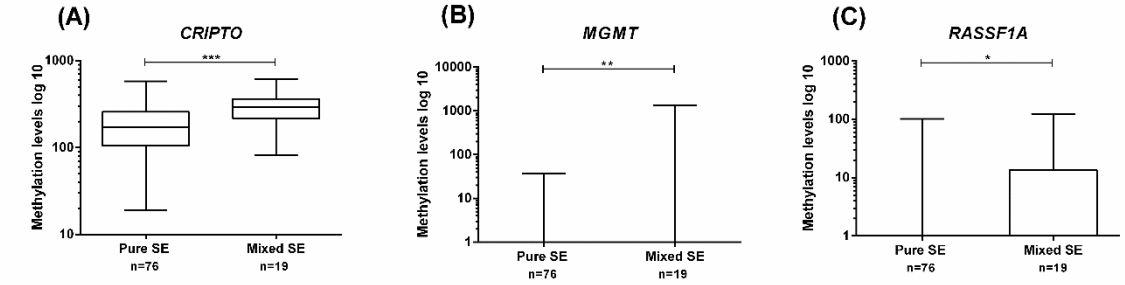


Figure 13. Boxplots of (A) *CRIPTO*, (B) *MGMT* and (C) *RASSF1A* promoter methylation levels in pure SE and SE as component of mixed TGCT samples. (Mann-Whitney U Test, *p<0.05, **p<0.01 and ***p<0.001).

3.2.2. Pure Embryonal Carcinoma *versus* Embryonal Carcinoma in Mixed Germ Cell Tumors

EmbrCa as component of mixed TGCT showed significantly higher *MGMT* methylation levels when compared with its pure form (p=0.0321) (Figure 14).

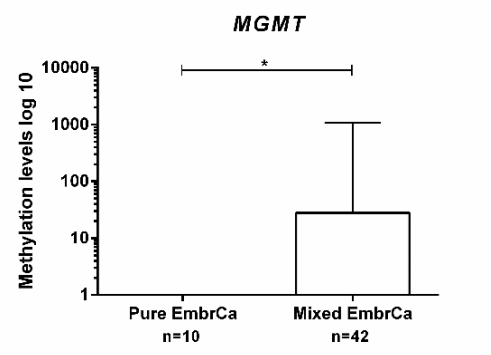


Figure 14. Boxplot of *MGMT* promoter methylation levels in pure EmbrCa and EmbrCa as component of mixed TGCT samples. (Mann-Whitney U Test, *p<0.05).

3.2.3. Pure Postpubertal-Type Teratoma *versus* Postpubertal-Type Teratoma in Mixed Germ Cell Tumors

HOXA9, *RASSF1A* and *SCGB3A1* methylation levels were significantly higher in postpubertal-type TE as component of mixed TGCT compared with pure postpubertal-type TE (p=0.0305, p=0.0140 and p=0.0099, respectively) (Figure 15).

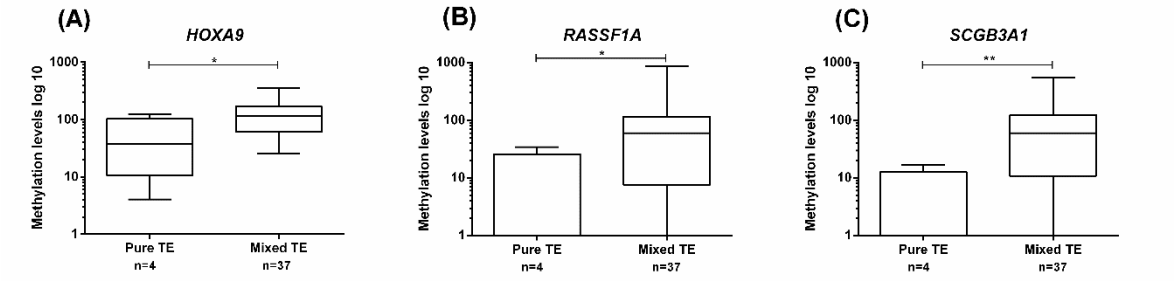


Figure 15. Boxplots of (A) *HOXA9*, (B) *RASSF1A* and (C) *SCGB3A1* promoter methylation levels in pure postpubertal-type TE and postpubertal-type TE as component of mixed TGCT samples. (Mann-Whitney U Test, *p<0.05, **p<0.01).

3.3. Comparisons Among Pure TGCT Subtypes

CRIPTO, *HOXA9*, *RASSF1A* and *SCGB3A1* methylation levels were significantly different among the several pure forms of GCNIS-related TGCT (Figure 16, Table 16).

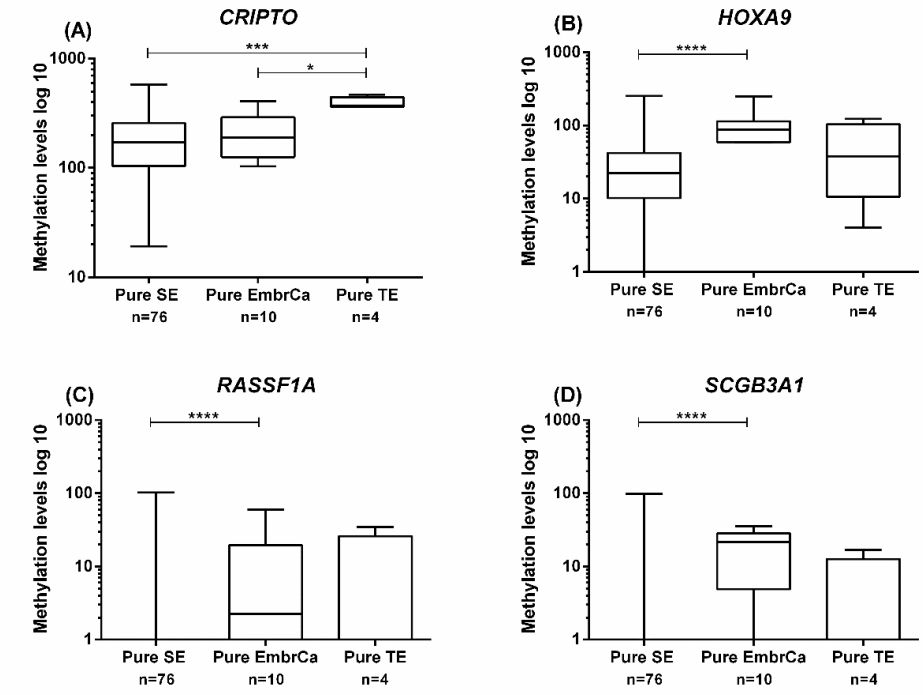


Figure 16. Boxplots of (A) *CRIPTO*, (B) *HOXA9*, (C) *RASSF1A* and (D) *SCGB3A1* promoter methylation levels in pure TGCT subtypes related to GCNIS. (TE: Postpubertal-type TE; Mann-Whitney U Test, *p<0.05, ***p<0.001, ****p<0.0001).

Indeed, *HOXA9*, *RASSF1A* and *SCGB3A1* methylation levels in SE differed significantly from EmbrCa, whilst *CRIPTO* methylation levels differed between SE and TE and between the latter and EmbrCa.

Table 16. Kruskal-Wallis and Mann-Whitney tests analysis of *CRIPTO*, *HOXA9*, *RASSF1A* and *SCGB3A1* promoter methylation levels between pure GCNIS-related TGCT samples.

PURE SUBTYPES COMBINATIONS	*CRIPTO*	*HOXA9*	*RASSF1A*	*SCGB3A1*
SE *vs*. EmbrCa	0.5982	**<0.0001**	**<0.0001**	**<0.0001**
SE *vs*. TE	**0.0009**	0.3392	0.0999	0.8486
EmbrCa *vs*. TE	**0.0140**	0.0999	0.2278	0.0509
p-value[1]	**0.0124**	**<0.0001**	**<0.0001**	**<0.0001**

SE: Seminomas; EmbrCa: Embryonal carcinoma; TE: Postpubertal-type teratoma;
p-values for all pair-wise comparisons were obtained using Mann-Whitney U Test with Bonferroni correction;
[1]: p-value Kruskal-Wallis analysis.

3.4. Comparisons Amongst the Components of Mixed Germ Cell Tumors

HOXA9, *MGMT*, *RASSF1A* and *SCGB3A1* methylation levels significantly differed amongst GCNIS-related TGCT subtypes as components of mixed TGCT (Figure 17, Table 17).

Overall, for all the analyzed genes, the lowest methylation levels were observed in mixed SE. Concerning *MGMT*, methylation levels did not differ among the different subtypes, except between mixed SE and TE. *RASSF1A* and *SCGB3A1* methylation levels followed a similar pattern among the components of mixed germ cell tumors. No significant differences in methylation levels were found between mixed YST, CH and TE for any of the analyzed genes.

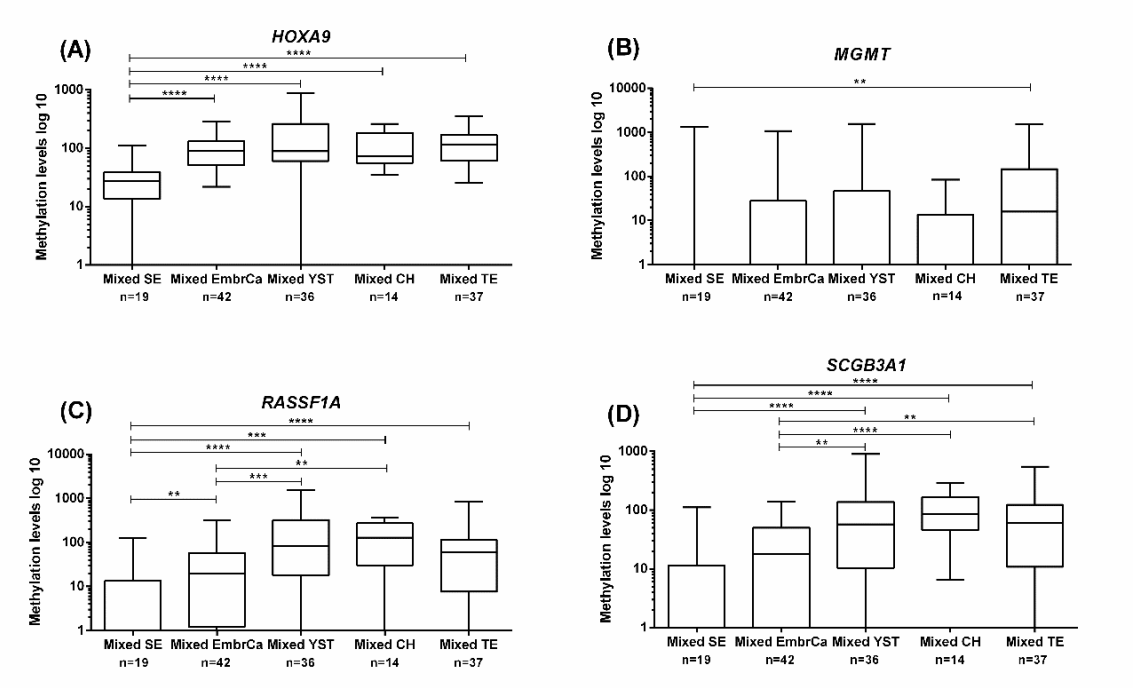


Figure 17. Boxplots of (A) *HOXA9*, (B) *MGMT*, (C) *RASSF1A* and (D) *SCGB3A1* promoter methylation levels in GCNIS-related TGCT subtypes as components of mixed TGCT samples. (YST: Postpubertal-type YST; TE: Postpubertal-type TE; Mann-Whitney U Test, **p<0.01;***p<0.001, ****p<0.0001).

Table 17. Kruskal-Wallis and Mann-Whitney tests analysis of *HOXA9*, *MGMT*, *RASSF1A* and *SCGB3A1* promoter methylation levels between GCNIS-related TGCT as components of mixed TGCT.

MIXED SUBTYPES COMBINATIONS	*HOXA9*	*MGMT*	*RASSF1A*	*SCGB3A1*
SE *vs.* EmbrCa	**<0.0001**	0.2093	**0.0047**	0.0320
SE *vs.* YST	**<0.0001**	0.1579	**<0.0001**	**<0.0001**
SE *vs.* CH	**<0.0001**	0.8396	**0.0002**	**<0.0001**
SE *vs.* TE	**<0.0001**	**0.0081**	**<0.0001**	**<0.0001**
EmbrCa *vs.* YST	0.2646	0.7523	**0.0007**	**0.0032**
EmbrCa *vs.* CH	0.7709	0.3711	**0.0048**	**<0.0001**
EmbrCa *vs.* TE	0.0937	0.0271	0.0161	**0.0022**
YST *vs.* CH	0.3679	0.2586	0.9744	0.2065
YST *vs.* TE	0.8277	0.0698	0.1801	0.9979
CH *vs.* TE	0.4383	0.0177	0.1436	0.1436
p-value[1]	**<0.0001**	**0.0201**	**<0.0001**	**<0.0001**

SE: Seminomas; EmbrCa: Embryonal carcinoma; YST: Postpubertal-type yolk-sac tumor; CH: Choriocarcinoma; TE: Postpubertal-type teratoma;
p-values for all pair-wise comparisons were obtained using Mann-Whitney U Test with Bonferroni correction;
[1]: p-value Kruskal-Wallis analysis.

3.5. Prepubertal-Type Subtypes *versus* Matching Postpubertal-Type Tumors

CRIPTO methylation levels were significantly higher in postpubertal-type YST from mixed TGCT compared to prepubertal-type YST (p=0.0086) (Figure 18), whereas the opposite was depicted for *HOXA9* (p=0.0368) (Figure 19).

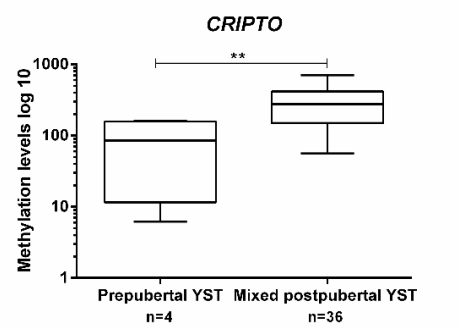


Figure 18. Boxplot of *CRIPTO* promoter methylation levels in prepubertal-type YST and postpubertal-type YST as component of mixed TGCT samples. (Mann-Whitney U Test, **p<0.01).

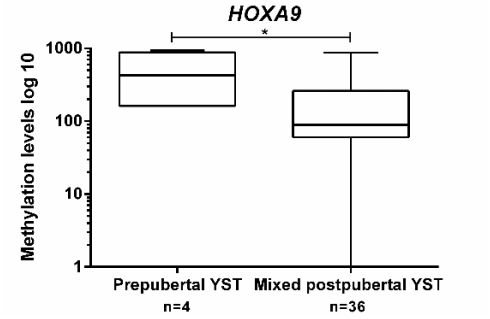


Figure 19. Boxplot of *HOXA9* promoter methylation levels in prepubertal-type YST and postpubertal-type YST as component of mixed TGCT samples. (Mann-Whitney U Test, *p<0.05).

Concerning TE, *HOXA9*, *MGMT* and *SCGB3A1* methylation levels were significantly higher in postpubertal-type TE as component of mixed TGCT compared to prepubertal-type TE (p=0.0091, p=0.0051 and p=0.0054, respectively) (Figure 20).

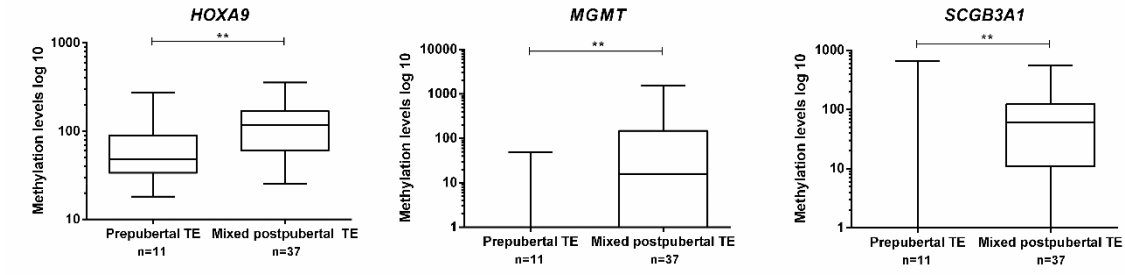


Figure 20. Boxplots of (A) *HOXA9*, (B) *MGMT* and (C) *SCGB3A1* promoter methylation levels in prepubertal-type TE and postpubertal-type TE as component of mixed TGCT samples. (Mann-Whitney U Test, **p<0.01).

4. ASSOCIATION BETWEEN GENE PROMOTER METHYLATION LEVELS AND CLINICOPATHOLOGICAL PARAMETERS

In postpubertal TGCT, a significant inverse correlation was found between methylation levels of all genes and patients' age (*CRIPTO*, *HOXA9*, *MGMT*, *RASSF1A* and *SCGB3A1*, p=0.027, r=-0.183; p<0.001, r=-0.356; p=0.001, r=-0.281; p<0.001, r=-0.375; p<0.001, r=-0.388 respectively). However, no correlation with age was found when pure SE and NST were considered separately. Contrarily, in prepubertal TGCT, a significant positive correlation was found between *CRIPTO* promoter methylation and patients' age (p=0.005, r=0.679). Moreover, no association was found between methylation levels and age of controls.

Considering the potential of five genes promoter methylation levels as diagnostic biomarkers, *CRIPTO* methylation levels displayed higher sensitivity (65.8%) for GCNIS-related TGCT detection compared to any of the classical serum tumor markers (AFP: 35.0%; β-hCG: 51.8% and LDH: 42.9%), assessed prior to orchiectomy (Figure 21).

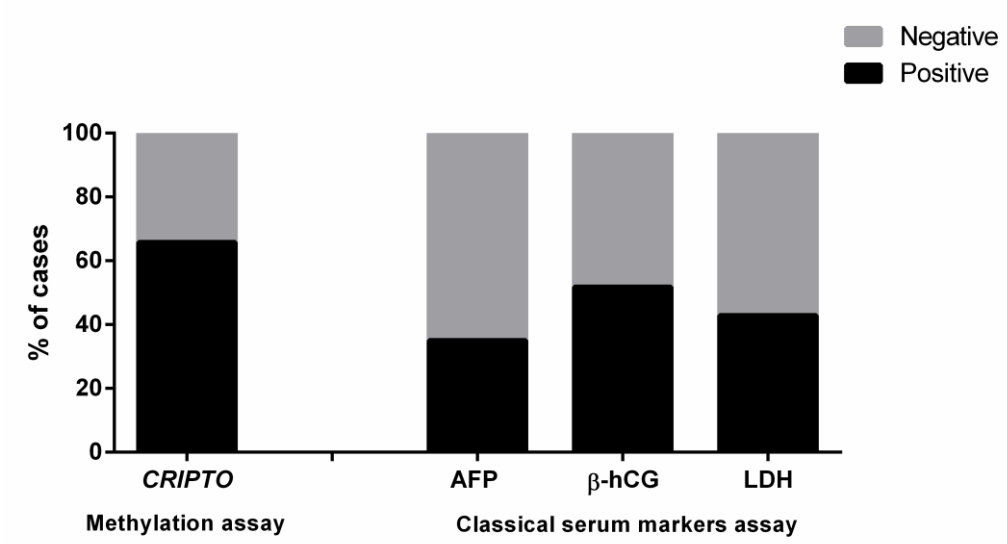


Figure 21. Graphic representation of sensitivity of *CRIPTO* methylation assay and classical serum tumor markers in TGCT related to GCNIS detection.

HOXA9, *RASSF1A* and *SCGB3A1* methylation levels significantly associated with advanced stage (stages II and III). Specifically, stage III patients displayed significantly higher methylation levels compared with stage I (*HOXA9*: p=0.0004, *RASSF1A*: p=0.0007 and *SCGB3A1*: p=0.0002) and stage II patients (*HOXA9*: p=0.0081, *RASSF1A*: p=0.0138 and *SCGB3A1*: p=0.0081) (Figure 22).

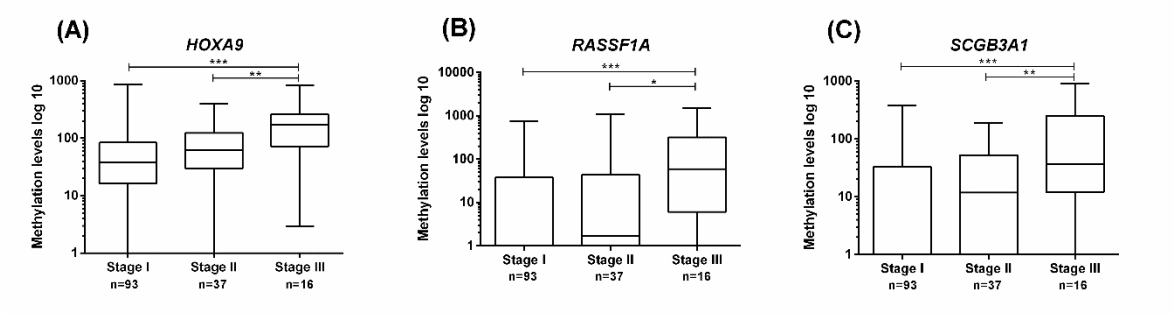


Figure 22. Boxplots of (A) *HOXA9*, (B) *RASSF1A* and (C) *SCGB3A1* promoter methylation levels in the different stages of the disease. (Mann-Whitney U Test, *p<0.05, **p<0.01; ***p<0.001).

Furthermore, *HOXA9*, *RASSF1A* and *SCGB3A1* methylation levels also significantly associated with prognostic-based staging. Indeed, higher methylation levels were depicted by patients classified in intermediate or poor prognosis compared to patients with good prognosis [*HOXA9*: p=0.0031 (good *vs.* poor), *RASSF1A*: p=0.0103 (good *vs.* intermediate) and *SCGB3A1*: p=0.0035 (good *vs.* poor)] (Figure 23). No significant associations were found for the remaining genes both for staging and prognostic-based staging.

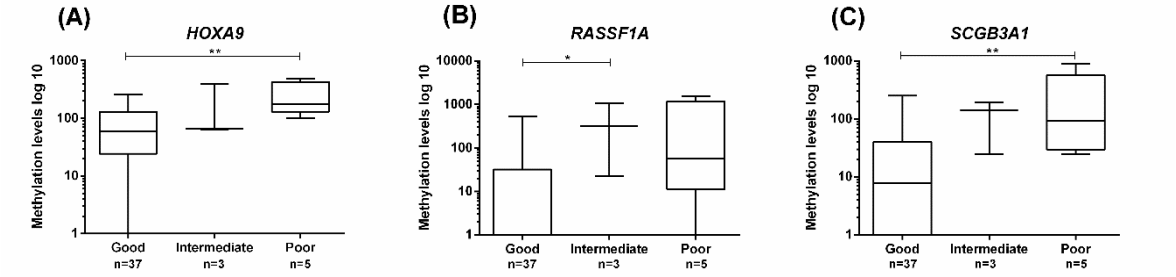


Figure 23. Boxplots of (A) *HOXA9*, (B) *RASSF1A* and (C) *SCGB3A1* promoter methylation levels in the different prognostic-based staging of the disease. (Mann-Whitney U Test, *p<0.05; **p<0.01).

V. DISCUSSION

Although TGCT account for only 1% of all male cancers worldwide, they are the most common cancer diagnosed in young Caucasian men [2] and represent a very heterogeneous group of tumors, although cytogenetically similar, suggesting that epigenetic alterations may contribute for this heterogeneity [121]. TGCT diagnosis is based on an extensive characterization of orchiectomy specimens, at morphological and immunohistochemical level, performed by trained pathologists. Correct differential diagnosis of TGCT is required for therapy decision and prognostication, which is often difficult due to TGCT histopathological variability [122]. Serum tumor markers levels measured before and after surgery provide additional information for diagnosis and patient monitoring. However, they display relatively low sensitivity and specificity for TGCT detection [77]. Thus, the identification and validation of new epigenetic-based biomarkers, displaying higher sensitivity and specificity, for screening, diagnosis and disease monitoring constitute of the main focus of TGCT research at present [77, 123]. The main goal of this study was to evaluate whether specific genes methylation profiles associated with distinct TGCT subtypes and may, thus, serve as type-specific biomarkers. Indeed, accurate TGCT subtype discrimination is essential, providing additional information to standard clinical, serological, imagiological and pathological parameters, impacting on staging and subsequent therapeutic decisions.

Based on assessment of aberrant DNA methylation at promoter regions, we firstly evaluated the diagnostic performance of a five genes panel. Considering GCNIS-related TGCT, we found that *CRIPTO*, *HOXA9*, *MGMT* and *SCGB3A1* methylation levels were significantly higher compared to controls. This observation contrasts with that of Brait *et al.* in which no significant differences between TGCT and control tissues were disclosed for *MGMT* methylation levels [124]. It should be emphasized, however, that our series more than quadruplicates that of Brait and co-workers (238 *vs.* 57), that might influence statistical analysis. Moreover, our control samples were collected from patients with testicular inflammatory disease or benign tumors, whereas those used by Brait *et al.* were obtained from patients submitted to surgical castration due to prostate cancer [124]. Nevertheless, despite the significant differences in methylation levels, ROC curve analysis disclosed a modest AUC value for each gene (not higher than 0.70), limiting its potential as biomarkers, individually.

Interestingly, although no significant differences in *RASSF1A* methylation levels were found between controls and GCNIS-related TGCT in general, significant *RASSF1A* promoter hypomethylation was observed in pure SE tissues. In this setting, the biomarker discriminated pure SE from controls with AUC of 0.8832, 87.5% sensitivity and 93.4% specificity. The performance of *RASSF1A* was, thus, superior to that of a previously reported gene panel [124]. On the other hand, *CRIPTO*, *HOXA9*, *MGMT* and *SCGB3A1*

promoter methylation levels were significantly higher in NST compared to controls, and a panel comprising *CRIPTO*/*HOXA9*/*SCGB3A1* achieved better performance than individual genes, although with modest specificity and NPV (68.8% and 73.3%, respectively). In the previously mentioned report, *APC*/*ER-β*/*MLH1* panel displayed AUC value of 0.89, with 85.7% sensitivity and 82.6% specificity for NST detection [124]. Nevertheless, direct comparisons with our study cannot be made because the composition of NST in Brait *et al.* series was not disclosed, the sample size is smaller and the same cut-off obtained by ROC curve analysis to discriminate between tumor and normal tissues was used for all comparisons [124]. Concerning methylation frequencies, the highest values for *HOXA9*, *RASSF1A* and *SCGB3A1* were depicted in CH as component of mixed germ cell tumors [7/14 (50%), 10/14 (71.4%) and 12/14 (86%), respectively], whilst the highest *HOXA9*, *MGMT* and *SCGB3A1* methylation frequencies were observed in postpubertal-type TE as component of mixed germ cell tumors [27/37 (73%), 21/37 (57%) and 22/37 (59.5%), respectively]. Lind *et al.* found that CH subtype displayed the highest *HOXA9*, *RASSF1A* and *SCGB3A1* methylation frequencies [119]. However, only one sample of pure CH was included in their study, while in our study all the CH samples (n=14) belonged to mixed TGCT. Furthermore, Koul *et al.* showed that the highest methylation frequencies of *RASSF1A* and *MGMT* were found in YST and EmbrCa, respectively [125]. Once more, comparisons among these studies should be made with precaution, as it should be recalled that in our study each component of mixed TGCT subtypes was analyzed separately, while in the aforementioned study, mixed TGCT were considered as a single subtype.

The main goal of this dissertation was to assess the methylation profile of some cancer-related genes among different TGCT subtypes. Altogether, GCNIS and SE have been characterized by global DNA hypomethylation, whereas NST generally disclose increased DNA methylation [37, 126]. Thus, a tumorigenic model suggests that SE arises from GCNIS cells that retain the methylation erasure that occurs during normal germ cell development, and subsequent *de novo* methylation is related to development of NST [37, 126]. Remarkably, our results are in accordance with this proposal, as *CRIPTO*, *HOXA9*, *MGMT*, *RASSF1A* and *SCGB3A1* promoter methylation were significantly more common in NST compared to pure SE, paralleling other previous observations, although frequencies vary among studies [119, 125-128]. These discrepancies might be due to differences in the number of samples analyzed in each study (our series is the largest), as well as of the TGCT subtypes, since we segregated pure NST from components of mixed TGCT subtypes. Moreover, different methodological approaches were used for methylation analysis, with most researchers using conventional qualitative MSP, whereas we used quantitative MSP, that displays higher sensitivity and specificity [129-131]. Notwithstanding, all these data suggest that a methylation panel could discriminate pure SE from NST. Accordingly, we

found that the *HOXA9*/*RASSF1A* panel displayed the best performance in this setting, with 90.1% sensitivity, 81.6% specificity and 87.4% accuracy. This result compares well with that of Brait *et al.*, which found that *APC*/*MLH1*/*ER-β*/*FKBP4* panel could discriminate pure SE from NST with 85.7% sensitivity and 93.0% specificity [124].

Contrarily to most previous studies aimed at characterizing altered gene promoter methylation patterns in TGCT, we decided to segregate “pure” forms of TGCT from those composing mixed TGCT. This enabled us to verify whether in morphologically and immunohistochemically overlapping lesions (*e.g.*, pure SE and SE foci in mixed TGCT) might exist differences in DNA methylation. Remarkably, we found that *CRIPTO*, *MGMT* and *RASSF1A* methylation levels were significantly higher in SE as component of mixed TGCT than in pure SE, and the same was found concerning *MGMT* methylation levels in EmbrCa as component of mixed TGCT compared to pure EmbrCa. A possible explanation is that in mixed TGCT, SE and EmbrCa cells may have already undergone molecular alterations that bring them closer to the other NST components compared to their pure form, as NST development may result from EmbrCa cell differentiation, that in turn may derive from direct reprogramming of GCNIS or SE cells [5]. Thus, although pure TGCT types and those composing mixed TGCT are histopathologically indistinguishable [42], they bear some differences in promoter methylation levels which could mark progression in this cancer model. Importantly, these findings clearly indicate that separate analyses for pure TGCT forms and for those collected from mixed TGCT are required, and that they should not be lumped together.

Variations in *CRIPTO* promoter methylation levels amongst TGCT subtypes might also be explained by the same tumorigenesis model. Indeed, we found that *CRIPTO* promoter methylation levels were significantly higher in pure postpubertal-type TE compared to pure SE and pure EmbrCa. These results are in accordance in previous observations by Spiller *et al.*, which found higher *CRIPTO* methylation levels in CH, TE and YST, correlating with low *CRIPTO* expression in CH and TE, and low *CRIPTO* methylation levels in EmbrCa and SE, correlating with high *CRIPTO* expression in EmbrCa [132]. Thus, *CRIPTO* expression seems to be associated with more undifferentiated TGCT (pure EmbrCa and pure SE), whereas *CRIPTO* promoter methylation characterizes more differentiated TGCT (pure postpubertal-type TE) [133].

TGCT unrelated to GCNIS - prepubertal-type YST and prepubertal-type TE - are a quite distinct group, although bearing similar designation to their postpubertal-type counterparts. Indeed, prepubertal-type YST morphologically overlaps postpubertal-type YST but it is less clinically aggressive, whereas prepubertal-type TE may resemble its postpubertal counterpart but they have quite distinct behavior, with the former behaving as a benign tumor and the latter as a malignant one [1]. Remarkably, we showed that, at molecular level,

CRIPTO and *HOXA9* methylation levels significantly differed between prepubertal-type YST and postpubertal-type YST as component of mixed TGCT, while *HOXA9*, *MGMT* and *SCGB3A1* methylation levels were also significantly different between prepubertal-type TE and postpubertal-type TE as component of mixed TGCT. Thus, it is tempting to speculate whether a different epigenetic landscape might contribute to rather dissimilar clinical behavior.

Another aim of this dissertation was to seek for associations between gene promoter methylation and standard clinical and pathological parameters. Interestingly, a significant negative correlation was found with patients' age. This result is probably due to the methylation profile found for specific TGCT subtypes and their age distribution. Indeed, our cohort included NST (n=70) and SE (n=76) with a median patients' age of 27 and 35 years, respectively, which is in accordance with the "bell-shaped" age distribution (NST: 25 years; SE: 35 years) reported for these two major TGCT subtypes [25]. Because SE displayed the lower gene promoter methylation levels, and affect older patients, a negative correlation between methylation levels and age could be expected. A similar explanation applies for the association between *CRIPTO* promoter methylation levels and age in prepubertal-type TGCT. Moreover, *HOXA9*, *RASSF1A* and *SCGB3A1* promoter methylation was significantly associated with advanced group staging and with prognostic-based staging for metastatic disease. All these genes displayed significantly higher promoter methylation levels in stage III disease. Furthermore, *HOXA9* and *SCGB3A1* methylation levels were significantly associated with poor prognosis groups and *RASSF1A* with intermediate prognosis groups. Interestingly, *CRIPTO* and *MGMT* promoter methylation did not associate with any clinicopathological parameter, which in accordance with previous publications [124, 134]. Altogether, these suggest that DNA methylation-based biomarkers might improve current risk stratification nomograms, especially for patients with metastatic disease, with a possible impact in therapeutic and monitoring strategy.

VI. CONCLUSIONS AND FUTURE PERSPECTIVES

In this study, we showed the potential of methylation-based biomarkers for TGCT detection, particularly NST detection, for which *CRIPTO*/*HOXA9*/*SCGB3A1* panel demonstrated an interesting performance. Moreover, we revealed promising methylation markers for detection of GCNIS-related TGCT. We confirmed differences in methylation patterns, previously reported for SE *vs.* NST, and we identified a two-gene (*HOXA9*/*RASSF1A*) panel that might discriminate between these two major TGCT subtypes with high sensitivity, specificity and accuracy. Furthermore, our study disclosed, for the first time, differences in specific gene promoter methylation levels between "pure" GCNIS-related TGCT and matching tumor foci from mixed GCNIS-related TGCT, demonstrating that despite morphological similarities, there are some differences at molecular level, which might be related with this quite unique model of tumorigenesis. Furthermore, differences in promoter methylation between pre- and postpubertal TGCT subtypes were also found. Finally, *HOXA9*, *RASSF1A* and *SCGB3A1* promoter methylation levels might provide useful information for risk stratification, improving current staging protocols, and might also have prognostic value.

In the future, we aim to evaluate the corresponding genes' expression in the same cohort to determine whether DNA promoter methylation is the main mechanism associated with regulation of gene expression among TGCT subtypes, serving as tool for genomic plasticity along tumorigenesis. Furthermore, we also intend to validate the genes-panel in liquid biopsies, aiming at providing more sensitive and specific biomarkers for TGCT detection and monitoring, eventually discriminating among subtypes, complementing standard clinical and pathological parameters. Because assessment of gene methylation status in cell-free plasma/serum DNA from patients with TGCT has already been reported as potential biomarker for TGCT diagnosis [135, 136], they might advantageously replace the classical serological tumor markers in routine clinical practice.

VII. REFERENCES

1. Moch, H., et al., *The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs—Part A: Renal, Penile, and Testicular Tumours.* European urology, 2016. **70**(1): p. 93-105.
2. Ferlay, J., et al., *Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012.* International journal of cancer, 2015. **136**(5): p. E359-E386.
3. Ferlay J, S.I., Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray, F. *GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer; 2013*. Available from: http://globocan.iarc.fr, 2013 [cited Accessed on 06/Januay/2017.
4. Trabert, B., et al., *International patterns and trends in testicular cancer incidence, overall and by histologic subtype, 1973–2007.* Andrology, 2015. **3**(1): p. 4-12.
5. Rijlaarsdam, M.A. and L.H. Looijenga. *An oncofetal and developmental perspective on testicular germ cell cancer*. in *Seminars in cancer biology*. 2014. Elsevier.
6. Chieffi, P., *Recent advances in molecular and cell biology of testicular germ cell tumors.* Int Rev Cell Mol Biol, 2014. **312**: p. 79-100.
7. Czene, K., P. Lichtenstein, and K. Hemminki, *Environmental and heritable causes of cancer among 9.6 million individuals in the Swedish family-cancer database.* International Journal of Cancer, 2002. **99**(2): p. 260-266.
8. Nathanson, K.L., et al., *The Y deletion gr/gr and susceptibility to testicular germ cell tumor.* The American Journal of Human Genetics, 2005. **77**(6): p. 1034-1043.
9. Linger, R., et al., *A physical analysis of the Y chromosome shows no additional deletions, other than Gr/Gr, associated with testicular germ cell tumour.* British journal of cancer, 2007. **96**(2): p. 357-361.
10. Kratz, C.P., P.L. Mai, and M.H. Greene, *Familial testicular germ cell tumours.* Best Practice & Research Clinical Endocrinology & Metabolism, 2010. **24**(3): p. 503-513.
11. Horvath, A., et al., *Functional phosphodiesterase 11A mutations may modify the risk of familial and bilateral testicular germ cell tumors.* Cancer research, 2009. **69**(13): p. 5301-5306.
12. Rapley, E.A., et al., *A genome-wide association study of testicular germ cell tumor.* Nature genetics, 2009. **41**(7): p. 807-810.
13. Hemminki, K. and X. Li, *Familial risk in testicular cancer as a clue to a heritable and environmental aetiology.* British journal of cancer, 2004. **90**(9): p. 1765-1770.
14. Hemminki, K. and B. Chen, *Familial risks in testicular cancer as aetiological clues.* International journal of andrology, 2006. **29**(1): p. 205-210.
15. Dieckmann, K.-P. and U. Pichlmeier, *Clinical epidemiology of testicular germ cell tumors.* World journal of urology, 2004. **22**(1): p. 2-14.
16. Wood, H.M. and J.S. Elder, *Cryptorchidism and testicular cancer: separating fact from fiction.* The Journal of urology, 2009. **181**(2): p. 452-461.
17. Doria-Rose, V.P., M.L. Biggs, and N.S. Weiss, *Subfertility and the risk of testicular germ cell tumors (United States).* Cancer Causes & Control, 2005. **16**(6): p. 651-656.
18. Raman, J.D., C.F. Nobert, and M. Goldstein, *Increased incidence of testicular cancer in men presenting with infertility and abnormal semen analysis.* The Journal of urology, 2005. **174**(5): p. 1819-1822.
19. Walsh, T.J., et al., *Increased risk of testicular germ cell cancer among infertile men.* Archives of internal medicine, 2009. **169**(4): p. 351-356.
20. Hotaling, J.M. and T.J. Walsh, *Male infertility: a risk factor for testicular cancer.* Nature Reviews Urology, 2009. **6**(10): p. 550-556.
21. Otite, U., et al., *Testicular microlithiasis: is it a benign condition with malignant potential?* European urology, 2001. **40**(5): p. 538-542.
22. Cutcutache, I., et al., *Exome-wide sequencing shows low mutation rates and identifies novel mutated genes in seminomas.* European urology, 2015. **68**(1): p. 77-83.

23. McIntyre, A., et al., *Genes, chromosomes and the development of testicular germ cell tumors of adolescents and adults.* Genes, Chromosomes and Cancer, 2008. **47**(7): p. 547-557.
24. McIntyre, A., et al., *Amplification and overexpression of the KIT gene is associated with progression in the seminoma subtype of testicular germ cell tumors of adolescents and adults.* Cancer research, 2005. **65**(18): p. 8085-8089.
25. Ulbright, T., et al., *WHO Classification of of Tumours of the Urinary System and Male Genital Organs*. 2016.
26. Kristensen, D.G., et al., *Epigenetic features of testicular germ cell tumours in relation to epigenetic characteristics of foetal germ cells.* International Journal of Developmental Biology, 2013. **57**(2-3-4): p. 309-317.
27. Oosterhuis, J.W. and L.H. Looijenga, *Testicular germ-cell tumours in a broader perspective.* Nature Reviews Cancer, 2005. **5**(3): p. 210-222.
28. Litchfield, K., et al., *The genomic landscape of testicular germ cell tumours: from susceptibility to treatment.* Nature Reviews Urology, 2016. **13**(7): p. 409-419.
29. Dieckmann, K.P., et al., *Testicular biopsy for early cancer detection–objectives, technique and controversies.* International journal of andrology, 2011. **34**(4pt2): p. e7-e13.
30. Dieckmann, K.-P., et al., *Diagnosis of contralateral testicular intraepithelial neoplasia (TIN) in patients with testicular germ cell cancer: systematic two-site biopsies are more sensitive than a single random biopsy.* European urology, 2007. **51**(1): p. 175-185.
31. RAJPERT-DE MEYTS, E., et al., *The emerging phenotype of the testicular carcinoma in situ germ cell.* Apmis, 2003. **111**(1): p. 267-279.
32. Hoei-Hansen, C.E., et al., *Transcription factor AP-2γ is a developmentally regulated marker of testicular carcinoma in situ and germ cell tumors.* Clinical Cancer Research, 2004. **10**(24): p. 8521-8530.
33. Nettersheim, D., et al., *NANOG promoter methylation and expression correlation during normal and malignant human germ cell development.* Epigenetics, 2011. **6**(1): p. 114-122.
34. Gillis, A., et al., *Expression and interdependencies of pluripotency factors LIN28, OCT3/4, NANOG and SOX2 in human testicular germ cells and tumours of the testis.* International journal of andrology, 2011. **34**(4pt2): p. e160-e174.
35. Sonne, S.B., et al., *Identity of M2A (D2-40) antigen and gp36 (Aggrus, T1A-2, podoplanin) in human developing testis, testicular carcinoma in situ and germ-cell tumours.* Virchows Archiv, 2006. **449**(2): p. 200-206.
36. van der Zwan, Y.G., et al., *Gonadal maldevelopment as risk factor for germ cell cancer: towards a clinical decision model.* European urology, 2015. **67**(4): p. 692-701.
37. Netto, G.J., et al., *Global DNA hypomethylation in intratubular germ cell neoplasia and seminoma, but not in nonseminomatous male germ cell tumors.* Modern Pathology, 2008. **21**(11): p. 1337-1344.
38. Almstrup, K., et al., *Carcinoma in situ testis displays permissive chromatin modifications similar to immature foetal germ cells.* British journal of cancer, 2010. **103**(8): p. 1269-1276.
39. Novotny, G.W., et al., *MicroRNA expression profiling of carcinoma in situ cells of the testis.* Endocrine-related cancer, 2012. **19**(3): p. 365-379.
40. Bahrami, A., J.Y. Ro, and A.G. Ayala, *An overview of testicular germ cell tumors.* Archives of pathology & laboratory medicine, 2007. **131**(8): p. 1267-1280.
41. Berney, D.M., et al., *Malignant germ cell tumours in the elderly: a histopathological review of 50 cases in men aged 60 years or over.* Modern Pathology, 2008. **21**(1): p. 54-59.
42. Cheng, L., B. Lyu, and L.M. Roth, *Perspectives on testicular germ cell neoplasms.* Human Pathology, 2017. **59**: p. 10-25.

43. Kao, C.-S., et al., *Testicular embryonal carcinoma: a morphologic study of 180 cases highlighting unusual and unemphasized aspects.* The American journal of surgical pathology, 2014. **38**(5): p. 689-697.
44. Rajpert-De Meyts, E., et al., *Testicular germ cell tumours.* The Lancet, 2016. **387**(10029): p. 1762-1774.
45. Sweeney, C.J., et al., *Results and outcome of retroperitoneal lymph node dissection for clinical stage I embryonal carcinoma–predominant testis cancer.* Journal of clinical oncology, 2000. **18**(2): p. 358-358.
46. Ulbright, T.M., *Germ cell tumors of the gonads: a selective review emphasizing problems in differential diagnosis, newly appreciated, and controversial issues.* Modern Pathology, 2005. **18**: p. S61-S79.
47. Sesterhenn, I.A. and C.J. Davis Jr, *Pathology of germ cell tumors of the testis.* Cancer Control, 2004. **11**(6): p. 374-387.
48. STEINBRONN, D.V., T.H. HICKS, and W.H. CARREL, *Mixed germ cell testis tumor in an 86-year-old man.* The Journal of urology, 1999. **162**(1): p. 161.
49. Kao, C.-S., et al., *Solid pattern yolk sac tumor: a morphologic and immunohistochemical study of 52 cases.* The American journal of surgical pathology, 2012. **36**(3): p. 360-367.
50. Cao, D., et al., *SALL4 is a novel diagnostic marker for testicular germ cell tumors.* The American journal of surgical pathology, 2009. **33**(7): p. 1065-1077.
51. Alvarado-Cabrero, I., N. Hernández-Toriz, and G.P. Paner, *Clinicopathologic analysis of choriocarcinoma as a pure or predominant component of germ cell tumor of the testis.* The American journal of surgical pathology, 2014. **38**(1): p. 111-118.
52. Humphrey, P.A., *Choriocarcinoma of the Testis.* The Journal of urology, 2014. **192**(3): p. 934-935.
53. Reilley, M.J. and L.C. Pagliaro, *Testicular choriocarcinoma: a rare variant that requires a unique treatment approach.* Current oncology reports, 2015. **17**(2): p. 1-9.
54. Howitt, B.E. and D.M. Berney, *Tumors of the Testis: Morphologic Features and Molecular Alterations.* Surgical pathology clinics, 2015. **8**(4): p. 687-716.
55. Tarrant, W.P., B.A. Czerniak, and C.C. Guo, *Relationship between primary and metastatic testicular germ cell tumors: a clinicopathologic analysis of 100 cases.* Human pathology, 2013. **44**(10): p. 2220-2226.
56. Daugaard, G., et al., *Surveillance for stage I nonseminoma testicular cancer: outcomes and long-term follow-up in a population-based cohort.* Journal of Clinical Oncology, 2014. **32**(34): p. 3817-3823.
57. Carrière, P., P. Baade, and L. Fritschi, *Population based incidence and age distribution of spermatocytic seminoma.* The Journal of urology, 2007. **178**(1): p. 125-128.
58. Chung, P.W., et al., *Spermatocytic seminoma: a review.* European urology, 2004. **45**(4): p. 495-498.
59. Williamson, S.R., et al., *The WHO 2016 Classification of Testicular Germ Cell Tumours: a Review and Update from the ISUP Testis Consultation Panel.* Histopathology, 2016.
60. Goriely, A., et al., *Activating mutations in FGFR3 and HRAS reveal a shared genetic origin for congenital disorders and testicular tumors.* Nature genetics, 2009. **41**(11): p. 1247-1252.
61. Looijenga, L.H., et al., *Genomic and expression profiling of human spermatocytic seminomas: primary spermatocyte as tumorigenic precursor and DMRT1 as candidate chromosome 9 gene.* Cancer Research, 2006. **66**(1): p. 290-302.
62. Aggarwal, N. and A.V. Parwani, *Spermatocytic seminoma.* Archives of pathology & laboratory medicine, 2009. **133**(12): p. 1985-1988.
63. Kristensen, D.G., et al., *Heterogeneity of chromatin modifications in testicular spermatocytic seminoma point toward an epigenetically unstable phenotype.* Cancer genetics, 2012. **205**(9): p. 425-431.

64. Taskinen, S., et al., *Testicular tumors in children and adolescents.* Journal of pediatric urology, 2008. **4**(2): p. 134-137.
65. Ulbright, T.M. and R.H. Young. *Testicular and paratesticular tumors and tumor-like lesions in the first 2 decades*. in *Seminars in diagnostic pathology*. 2014. Elsevier.
66. Cornejo, K.M., et al., *Yolk sac tumor of the testis in infants and children: a clinicopathologic analysis of 33 cases.* The American journal of surgical pathology, 2015. **39**(8): p. 1121-1131.
67. Lee, S.D., *Epidemiological and clinical behavior of prepubertal testicular tumors in Korea.* The Journal of urology, 2004. **172**(2): p. 674-678.
68. Horwich, A., D. Nicol, and R. Huddart, *Testicular germ cell tumours.* Bmj, 2013. **347**(f5526): p. 1-9.
69. Albers, P., et al., *Guidelines on testicular cancer: 2015 update.* European urology, 2015. **68**(6): p. 1054-1068.
70. Dieckmann, K., et al., *MicroRNAs miR-371-3 in serum as diagnostic tools in the management of testicular germ cell tumours.* British journal of cancer, 2012. **107**(10): p. 1754-1760.
71. Gilligan, T.D., et al., *American Society of Clinical Oncology Clinical Practice Guideline on uses of serum tumor markers in adult males with germ cell tumors.* Journal of Clinical Oncology, 2010. **28**(20): p. 3388-3404.
72. Stevenson, S.M. and W.T. Lowrance, *Epidemiology and diagnosis of testis cancer.* Urologic Clinics of North America, 2015. **42**(3): p. 269-275.
73. SB, E., et al., *AJCC cancer staging manual, 7th edition*. 2010: Springer-Verlag New York.
74. Powles, T.B., et al., *The changing presentation of germ cell tumours of the testis between 1983 and 2002.* BJU international, 2005. **95**(9): p. 1197-1200.
75. DeSantis, C.E., et al., *Cancer treatment and survivorship statistics, 2014.* CA: a cancer journal for clinicians, 2014. **64**(4): p. 252-271.
76. Oldenburg, J., J.M. Martin, and S.D. Fosså, *Late relapses of germ cell malignancies: incidence, management, and prognosis.* Journal of Clinical Oncology, 2006. **24**(35): p. 5503-5511.
77. Murray, M.J., R.A. Huddart, and N. Coleman, *The present and future of serum diagnostic tests for testicular germ cell tumours.* Nature Reviews Urology, 2016.
78. Mead, G. and S. Stenning, *The International Germ Cell Consensus Classification: a new prognostic factor-based staging classification for metastatic germ cell tumours.* Clinical Oncology, 1997. **9**(4): p. 207-209.
79. Tandstad, T., et al., *Management of clinical stage I seminomatous testicular cancer: A report from SWENOTECA.* Journal of Clinical Oncology, 2014. **32**(15_suppl): p. 4508-4508.
80. Hanna, N.H. and L.H. Einhorn, *Testicular cancer—discoveries and updates.* New England Journal of Medicine, 2014. **371**(21): p. 2005-2016.
81. Albers, P., et al., *Risk factors for relapse in clinical stage I nonseminomatous testicular germ cell tumors: results of the German Testicular Cancer Study Group Trial.* Journal of clinical oncology, 2003. **21**(8): p. 1505-1512.
82. Dupont, C., D.R. Armant, and C.A. Brenner. *Epigenetics: definition, mechanisms and clinical perspective*. in *Seminars in reproductive medicine*. 2009. © Thieme Medical Publishers.
83. Deans, C. and K.A. Maggert, *What do you mean,"epigenetic"?* Genetics, 2015. **199**(4): p. 887-896.
84. Valdes-Mora, F. and S. Clark, *Prostate cancer epigenetic biomarkers: next-generation technologies.* Oncogene, 2015. **34**(13): p. 1609-1618.
85. Paluch, B.E., et al., *Epigenetics: A primer for clinicians.* Blood reviews, 2016. **30**(4): p. 285-295.
86. Sharma, S., T.K. Kelly, and P.A. Jones, *Epigenetics in cancer.* Carcinogenesis, 2010. **31**(1): p. 27-36.

87. Tsai, H.-C. and S.B. Baylin, *Cancer epigenetics: linking basic biology to clinical medicine.* Cell research, 2011. **21**(3): p. 502-517.
88. Esteller, M., *Non-coding RNAs in human disease.* Nature Reviews Genetics, 2011. **12**(12): p. 861-874.
89. Rounge, T., et al., *Profiling of the small RNA populations in human testicular germ cell tumors shows global loss of piRNAs.* Molecular cancer, 2015. **14**(1): p. 153.
90. Kaikkonen, M.U., M.T. Lam, and C.K. Glass, *Non-coding RNAs as regulators of gene expression and epigenetics.* Cardiovascular research, 2011. **90**(3): p. 430-440.
91. Mendell, J.T., *MicroRNAs: critical regulators of development, cellular physiology and malignancy.* Cell cycle, 2005. **4**(9): p. 1179-1184.
92. He, L. and G.J. Hannon, *MicroRNAs: small RNAs with a big role in gene regulation.* Nature Reviews Genetics, 2004. **5**(7): p. 522-531.
93. Garzon, R., G.A. Calin, and C.M. Croce, *MicroRNAs in cancer.* Annual review of medicine, 2009. **60**: p. 167-179.
94. Ellinger, J., S.C. Müller, and D. Dietrich, *Epigenetic biomarkers in the blood of patients with urological malignancies.* Expert review of molecular diagnostics, 2015. **15**(4): p. 505-516.
95. Liu, L., et al., *MicroRNA-302a sensitizes testicular embryonal carcinoma cells to cisplatin-induced cell death.* Journal of cellular physiology, 2013. **228**(12): p. 2294-2304.
96. Kumar, R., et al., *Epigenomic regulation of oncogenesis by chromatin remodeling.* Oncogene, 2016. **35**(34): p. 4423-4436.
97. Füllgrabe, J., E. Kavanagh, and B. Joseph, *Histone onco-modifications.* Oncogene, 2011. **30**(31): p. 3391-3403.
98. Kouzarides, T., *Chromatin modifications and their function.* Cell, 2007. **128**(4): p. 693-705.
99. Skene, P.J. and S. Henikoff, *Histone variants in pluripotency and disease.* Development, 2013. **140**(12): p. 2513-2524.
100. Zink, L.-M. and S.B. Hake, *Histone variants: nuclear function and disease.* Current opinion in genetics & development, 2016. **37**: p. 82-89.
101. Goldberg, A.D., C.D. Allis, and E. Bernstein, *Epigenetics: a landscape takes shape.* Cell, 2007. **128**(4): p. 635-638.
102. Esteller, M., *Epigenetics in cancer.* New England Journal of Medicine, 2008. **358**(11): p. 1148-1159.
103. De Carvalho, D.D., J.S. You, and P.A. Jones, *DNA methylation and cellular reprogramming.* Trends in cell biology, 2010. **20**(10): p. 609-617.
104. Tost, J., *DNA methylation: an introduction to the biology and the disease-associated changes of a promising biomarker.* Molecular biotechnology, 2010. **44**(1): p. 71-81.
105. Portela, A. and M. Esteller, *Epigenetic modifications and human disease.* Nature biotechnology, 2010. **28**(10): p. 1057-1068.
106. Jones, P.A. and G. Liang, *Rethinking how DNA methylation patterns are maintained.* Nature Reviews Genetics, 2009. **10**(11): p. 805-811.
107. Lo, R. and R. Weksberg, *Biological and biochemical modulation of DNA methylation.* 2014.
108. Kanwal, R. and S. Gupta, *Epigenetic modifications in cancer.* Clinical genetics, 2012. **81**(4): p. 303-311.
109. Brait, M. and D. Sidransky, *Cancer epigenetics: above and beyond.* Toxicology mechanisms and methods, 2011. **21**(4): p. 275-288.
110. Costa, A.L., et al., *The epigenetics of testicular germ cell tumors: looking for novel disease biomarkers.* Epigenomics, 2017. **9**(2): p. 155-169.
111. Bianco, C., et al., *Role of Cripto-1 in stem cell maintenance and malignant progression.* The American journal of pathology, 2010. **177**(2): p. 532-540.
112. Rodrigues, M.F.S.D., et al., *Methylation status of homeobox genes in common human cancers.* Genomics, 2016. **108**(5): p. 185-193.

113. Jacinto, F.V. and M. Esteller, *MGMT hypermethylation: a prognostic foe, a predictive friend.* DNA repair, 2007. **6**(8): p. 1155-1160.
114. Donninger, H., M.D. Vos, and G.J. Clark, *The RASSF1A tumor suppressor.* Journal of cell science, 2007. **120**(18): p. 3163-3172.
115. Krop, I., et al., *HIN-1, an inhibitor of cell growth, invasion, and AKT activation.* Cancer research, 2005. **65**(21): p. 9659-9669.
116. Cottrell, S.E. and P.W. Laird, *Sensitive detection of DNA methylation.* Annals of the New York Academy of Sciences, 2003. **983**(1): p. 120-130.
117. Eads, C.A., et al., *Epigenetic patterns in the progression of esophageal adenocarcinoma.* Cancer research, 2001. **61**(8): p. 3410-3418.
118. Ramalho-Carvalho, J., et al., *Altered expression of MGMT in high-grade gliomas results from the combined effect of epigenetic and genetic aberrations.* PLoS One, 2013. **8**(3): p. e58206.
119. Lind, G., et al., *Novel epigenetically deregulated genes in testicular cancer include homeobox genes and SCGB3A1 (HIN-1).* The Journal of pathology, 2006. **210**(4): p. 441-449.
120. Arya, M., et al., *Basic principles of real-time quantitative PCR.* Expert review of molecular diagnostics, 2005. **5**(2): p. 209-219.
121. Feldman, D.R., *Update in germ cell tumours.* Current opinion in oncology, 2015. **27**(3): p. 177-184.
122. Sharma, P., et al., *Disparities in Interpretation of Primary Testicular Germ Cell Tumor Pathology.* American journal of clinical pathology, 2015. **144**(2): p. 289-294.
123. Murray, M.J. and N. Coleman, *Testicular cancer: a new generation of biomarkers for malignant germ cell tumours.* Nature Reviews Urology, 2012. **9**(6): p. 298-300.
124. Brait, M., et al., *DNA methylation profiles delineate epigenetic heterogeneity in seminoma and non-seminoma.* British journal of cancer, 2012. **106**(2): p. 414-423.
125. Koul, S., et al., *Characteristic promoter hypermethylation signatures in male germ cell tumors.* Molecular cancer, 2002. **1**(1): p. 8.
126. Smiraglia, D.J., et al., *Distinct epigenetic phenotypes in seminomatous and nonseminomatous testicular germ cell tumors.* Oncogene, 2002. **21**(24): p. 3909.
127. Smith-Sørensen, B., et al., *Frequent promoter hypermethylation of the O6-Methylguanine-DNA Methyltransferase (MGMT) gene in testicular cancer.* Oncogene, 2002. **21**(57): p. 8878.
128. Honorio, S., et al., *Frequent epigenetic inactivation of the RASSF1A tumour suppressor gene in testicular tumours and distinct methylation profiles of seminoma and nonseminoma testicular germ cell tumours.* Oncogene, 2003. **22**(3): p. 461-466.
129. Hoque, M.O., et al., *Quantitative methylation-specific polymerase chain reaction gene patterns in urine sediment distinguish prostate cancer patients from control subjects.* Journal of Clinical Oncology, 2005. **23**(27): p. 6569-6575.
130. Fackler, M.J., et al., *Quantitative multiplex methylation-specific PCR assay for the detection of promoter hypermethylation in multiple genes in breast cancer.* Cancer research, 2004. **64**(13): p. 4442-4452.
131. Shames, D.S., J.D. Minna, and A.F. Gazdar, *Methods for detecting DNA methylation in tumors: from bench to bedside.* Cancer letters, 2007. **251**(2): p. 187-198.
132. Spiller, C.M., et al., *Cripto: Expression, epigenetic regulation and potential diagnostic use in testicular germ cell tumors.* Molecular oncology, 2016. **10**(4): p. 526-537.
133. Spiller, C.M., J. Bowles, and P. Koopman, *Nodal/Cripto signaling in fetal male germ cell development: implications for testicular germ cell tumors.* International Journal of Developmental Biology, 2013. **57**(2-3-4): p. 211-219.
134. da Silva Martinelli, C.M., A. van Helvoort Lengert, and F. Mavignier, *MGMT and CALCA promoter methylation are associated with poor prognosis in testicular germ cell tumor patients.* Oncotarget, 2016. **5**.
135. Kawakami, T., et al., *XISTunmethylated DNA fragments in male-derived plasma as a tumour marker for testicular cancer.* The Lancet, 2004. **363**(9402): p. 40-42.

136. Ellinger, J., et al., *CpG island hypermethylation of cell-free circulating serum DNA in patients with testicular cancer.* The Journal of urology, 2009. **182**(1): p. 324-329.

VIII. SUPPLEMENTARY MATERIAL

ANNEX I

The epigenetics of testicular germ cell tumors: looking for novel disease biomarkers

Abstract

Testicular germ cell tumors (TGCT) are a group of heterogeneous, biologically diverse and clinically challenging neoplasms. Despite the relatively low incidence and mortality rates, a subgroup of patients with disseminated disease will relapse after conventional therapy and have a dismal prognosis. Moreover, TGCT afflict mostly young men and have therapeutic peculiarities, with some patients showing resistance to cisplatin-based treatments and others being troubled by irreversible side effects, such as infertility. Most TGCT share a common tumorigenic pathway and are cytogenetically similar, making room for Epigenetics to explain its heterogeneity at pathological and clinical level. In this review, we summarize the foremost epigenetic alterations among TGCT focusing on their clinical potential as diagnostic, prognostic and predictive biomarkers.

Keywords

TGCT; GCNIS; DNA methylation; Histone modifications; miRNAs; piRNAs; biomarkers

Introduction

Despite their rarity, testicular germ cell tumors (TGCT) are the most common cancers in Caucasian men between 15 and 44 years in developed countries, accounting for only 1% of all male cancers worldwide, and constitute up to 60% of all malignancies diagnosed in men aged 20 to 40 years [1]. Besides representing over 95% of testicular neoplasms, germ cell tumors also comprise a small proportion of ovarian malignancies, and may even occur at extragonadal sites, mainly along the anatomical midline of the body, a fact related to embryogenesis [2]. The overall incidence (1.5 cases per 100.000) has increased over the last decades and 65.827 new cases of testicular cancer are expected worldwide in 2030 (i.e., 10.561 more than that in 2012) [1, 3]. According to the most recent data, 1 out of every 250 men living in the United States will be diagnosed with testicular cancer during their lifetime and out of the 8.720 new cases estimated for 2016, 380 will die of disease [4]. Remarkably, incidence rates show striking geographical variation, with larger increments observed in men living in developed countries, which is consistent with the suggested influence of environmental risk factors typical of Western lifestyle, in addition to the well-known impact of genetic susceptibility. Early *in utero* exposure to industrialized lifestyle factors that cause undervirilization of the embryo is also in line with the particular predilection of these tumors for young men [5]. In fact, the age distribution is approximately "bell-shaped" (median age of 35 years-old for seminomas (SE), 25 for non-seminomatous tumors (NST) and 30 for mixed tumors), which is presumably explained by the superior aggressiveness of NST that entails earlier clinical manifestation.

Clinical relevance

Opposed to the crescent incidence, TGCT mortality has been decreasing over the past decades, mostly due to effective multimodal treatment. TGCT are among the most curable solid neoplasms, with 5-year survival rates over 95%. Even when widely metastasized, survival rates are over 80%, reflecting general sensitivity to DNA-damaging agents, particularly for SE [4, 6]. This sharply contrasts with the clinical scenario five decades ago, when diagnosis of disseminated TGCT meant death within one year for about 90% of patients [7]. Putting together both the paucity and good prognosis of TGCT one might argue whether further research efforts in this field are still necessary. Nevertheless, there are many reasons for TGCT to remain on the brink: despite outstanding cure rates, about 15-20% of patients with disseminated disease will relapse, and

late relapses (>2 years after remission), in particular, display poor prognosis; some tumors are resistant to cisplatin and the genetic and epigenetic mechanisms underlying this resistance need further clarification; in a cancer with such high cure rates, an attempt to reduce morbidity is mandatory, since chemotherapy and radiotherapy entail important side effects in young patients (secondary cancers, cardiovascular disease, etc) whose real extent is not yet known and that could shift the optimal management to active surveillance; as these tumors affect mostly young men, discussion of reproductive health issues is mandatory, namely sperm cryoconservation, testosterone supplementation and contraception during chemotherapy; and apart from the pathognomonic gain of chromosome 12p [mainly as isochromosome 12p (i12p)], and despite recent efforts to discover other genetic/epigenetic foci of susceptibility, no validated molecular markers exist that may be used for screening, diagnostic or prognostic purposes [2, 8-13].

Pathological characterization of TGCT in brief

TGCT are very heterogeneous neoplasms, probably owing to their origin from pluripotent neoplastic germ cells. Natural consequences of this complexity include the diagnostic challenges and dilemmas imposed to Pathologists, justifying the various classification schemes created throughout the years. Correct diagnosis and staging of TGCT require extensive and systematic sampling of surgical specimens of orchiectomy, as well as detailed morphological and immunohistochemical analysis to identify the presence and amount of different components, which has profound implications on staging and subsequent therapeutic strategies. Moreover, an extensive search for germ cell neoplasia *in situ* (GCNIS) in the surrounding testicular parenchyma should be sought to allow for correct categorization of TGCT, especially in cases of challenging differential diagnosis [14-18].

Due to TGCT extreme morphological heterogeneity and existence of many unusual patterns, differential diagnosis is challenging (Figure 1). SE is the most common type of TGCT, representing about 50% of cases, typically displaying a creamy, bulging cut surface and an arrangement in nests/sheets of cells separated by fibrous septa containing prominent lymphovascular infiltrate, in clear contrast with spermatocytic tumor (ST), a major differential diagnosis. In difficult cases, ST negativity for classic seminoma immunohistochemical markers is decisive. Concerning yolk-sac tumor (YST), few differences exist between prepubertal and postpubertal subtypes at the morphological level. YST may be incredibly heterogeneous, a feature that might

impair their recognition in mixed tumors. Nevertheless, reticular and microcystic patterns are the most frequent, and a myxoid stroma is commonly seen. Postpubertal teratoma (TE) usually has a heterogeneous cut surface and is composed of elements derived from ectodermal, mesodermal and/or endodermal elements. Immature components are frequently admixed with mature tissue but this finding has no prognostic significance, in clear contrast with its ovarian counterpart. Conversely, prepubertal TE includes epidermoid and dermoid cysts which bear close resemblance to ovarian mature teratomas. Embryonal carcinoma (EmbrCa) usually shows a predominantly solid pattern, with remarkable cytological atypia and extensive necrosis, whereas choriocarcinoma (CH) is an aggressive tumor (most patients presenting with metastatic disease), associated with extensive areas of hemorrhage and necrosis [14].

Cancer model and recent changes in classification

TGCT are presently separated into two major categories: those related and those unrelated to GCNIS [19]. This classification scheme has been recently adopted in the 4th Edition of the *World Health Organization Classification of Tumours of the Urinary System and Male Genital Organs* (released in 2016), shifting from an essentially morphological classification, which grouped very distinct tumors under very similar diagnostic categories, to a pathogenesis-based classification [20].

GCNIS derives from primordial germ cells/gonocytes that fail to differentiate into prespermatogonia and thus share many properties with these cells. GCNIS is found in the vicinity of most TGCT (both SE and NST) and virtually always progresses to overt cancer. It is also found in the contralateral testis of men previously diagnosed with TGCT in approximately 5% of cases. Polyploidization constitutes the initiation step for GCNIS formation, which only gains invasive potential after the hormonal changes that occur at puberty, progressing then to the default pathway and originating SE, or regaining pluripotency (a phenomenon known as “reprogramming)” and originating EmbrCa cells. These might then differentiate into the typical NST subtypes: embryonal differentiation originates EmbrCa; extraembryonal differentiation originates postpubertal-type YST and CH; and somatic differentiation originates postpubertal-type TE. This differentiation process is associated with *POU5F1* (*OCT3/4*) downregulation, a marker constitutively expressed in GCNIS, SE and EmbrCa. Indeed, GCNIS-derived neoplasms share many characteristics that

unite them in an individual category: similar epidemiology (postpubertal tumors, the most frequent TGCT, etc), association with perturbed testicular development, i12p and malignant behavior [21].

GCNIS-unrelated tumors are substantially less frequent and include prepubertal YST, prepubertal TE, and ST. The former two derive from a more immature primordial germ cell and occur in young children, while the latter originates from more mature spermatogonia/spermatocytes and typically afflicts adults aged more than 50 years. No i12p amplification is detected in these tumors and they mostly follow a clinically benign behavior [2, 20, 22].

Epigenetic aberrations in TGCT and their potential as diagnostic, prognostic & predictive biomarkers

TGCT represent a unique and complex cancer model and despite considerable progress in patient management, a complete understanding of all genetic and epigenetic alterations implicated in their genesis, which may help identify patients benefiting from additional therapeutic strategies or those that might avoid unnecessary treatment, is still an unmet goal [23]. Despite their heterogeneity, most TGCT share a common cytogenetic background and, consequently, epigenetic alterations may help illuminate the causes of phenotypic and clinical diversity. These alterations might constitute tumor-specific characteristics that may add relevant information to standard clinical, serological, imagiological and pathological parameters, constituting a novel type of TGCT biomarkers. Indeed, only 60% of TGCT patients exhibit increased levels of classical serum tumor markers human chorionic gonadotropin (HCG), α-fetoprotein (AFP), lactate dehydrogenase (LDH) or alkaline phosphatase. Thus, discovery and validation of new biomarkers is mandatory [24, 25], which should demonstrate higher sensitivity and specificity for diagnosis and disease monitoring than the existing ones [26]. For TGCT, the ideal biomarker should be stable in body fluids, exist at significantly higher levels in TGCT patients compared to healthy individuals and be detectable regardless of the age of the patient, anatomical site of disease or histological subtype [27, 28].

This review will focus on the foremost epigenetic alterations specifically involved in genesis and progression of TGCT, emphasizing their potential as novel cancer biomarkers to assist in clinical management of TGCT patients (Tables 1 & 2).

DNA methylation

Generally, cancer cells are characterized by global hypomethylation and locus-specific hypermethylation [29]. Interestingly, development of GCNIS has been associated with the period in which the under-methylated genome of primordial germ cells becomes heavily methylated [30]. Overall, both SE and GCNIS are mostly characterized by global hypomethylation, whereas NST exhibit methylation patterns more similar to those observed in other cancers derived from somatic cells [30-32]. Thus, promoter methylation status might allow for discrimination of TGCT from somatic cancers, which is extremely valuable for clinical management, as the differential diagnosis of testicular masses encompasses other lesions (both metastatic and even non-neoplastic, such as cysts and inflammatory lesions) and the testicle is seldom subjected to diagnostic biopsies given the risks of the procedure. Importantly, because SE and NST have different treatment modalities and prognosis, promoter methylation status may assist in the discrimination between these two TGCT subtypes [33].

The *XIST* gene represents a good example, as its CpG sites are methylated in male somatic cells whereas 5' regions are hypomethylated in TGCTs, regardless of gene expression [34]. Indeed, unmethylated *XIST* sequences were detected in 30/31 tissues (SE and NST) and in 16/25 plasma samples from patients with TGCT, contrasting with peripheral blood lymphocytes from healthy males (0/14) or plasma from patients with non-germ cells tumors (0/24) [34]. Thus, assessment of *XIST* unmethylated DNA fragments in plasma was suggested as a possible biomarker for detection and monitoring of TGCT patients [34]. However, these results need confirmation in larger series and attempts to increase sensitivity (e.g., using techniques such as real-time PCR [RT-PCR]) are expected to be looked up in the future.

Concerning gene-specific promoter hypermethylation, several examples have been reported in TGCT. Koul *et al.* evaluated the methylation status of 21 gene promoters in TGCT tissues and cell lines and most of them exhibited hypermethylation, although it was not found in controls (normal tissue samples) [33]. Interestingly, prevalence of gene promoter hypermethylation was significantly more pronounced in NST than in SE (38/63 NST vs. 5/29 SE), with the highest frequency depicted in YST, which might be of value for identification of an YST component among mixed TGCT, that may be overlooked. Promoter hypermethylation was associated with transcriptional silencing or downregulation of most genes (in all tumors for *MGMT* and *MLH1*, in 8/10 cases for *RASSF1A* and 3/5 cases for *RARB*), which was restored

after treatment with 5-aza-2' deoxycytidine (5-aza-dC), a widely used demethylating agent [33]. *MGMT* encodes for O6-methylguanine-DNA methyltransferase, that under normal conditions protects cells from the promutagenic O6-methyguanine alkylation by removing DNA adducts formed by alkylating agents [35]. *MGMT* promoter methylation was observed in TGCT tissues (32/69, 46.4%), contrasting with normal testis samples, and at a significantly higher rate in NST (24/35, 68.6%) compared to SE (8/33, 24.2%). Moreover, downregulated *MGMT* expression in six out of seven methylated TGCTs supports an association between altered gene expression and development of NST [36]. These results were recently confirmed in another study which disclosed that *MGMT* and *CALCA* were frequently methylated in NST (*MGMT*: 20/22, 90.9%, p=0.019; *CALCA*: 19/21, 90.5%, p=0.026) and associated with poor clinical outcome in TGCT patients (*MGMT*: 5-year event-free survival (EFS) rate of 50.5% for promoter methylation vs. 77.1% for unmethylated promoter, p=0.032; *CALCA*: 5-year EFS rate of 51.3% for promoter methylation vs. 77.0% for unmethylated promoter, p=0.029) [37]. *CALCA* methylation was also found to be significantly more common among EmbrCa and YST subtypes (p=0.017). Furthermore, *CALCA* promoter methylation was also significantly associated with disease refractory to therapy (9/19, 47.4%, p=0.005 of patients with methylated promoter vs. 6/19, 14.0%, p=0.005 of patients with unmethylated promoter status) [37]. Overall, these results suggest that gene-specific methylation status may be clinically relevant for prognostication of TGCT patients. Further studies on these markers should also be pursued as new treatment modalities targeting DNA methylation may be foreseen for refractory TGCT. Nevertheless, no clinical trials testing DNA methylation inhibitors in TGCT have been, thus far, conducted.

More genes have been shown to harbor promoter hypermethylation in NST, including *SCGB3A1* / *HIN-1* (19/35 in NST vs. 0/20 in SE), *RASSF1A* (10/35 in NST vs. 0/19 in SE) and *HOXA9* (9/35 in NST vs. 0/20 in SE) [38], with the lowest levels found in EmbrCa and the highest in TE. This epigenetic dysregulation may thus play an important role in TGCT development [38] and might be related with the extent of cell differentiation, with progressive increase of methylation along that process. EmbrCa is the most common NST subtype, which may be found pure or admixed with other NST. Global methylation analysis of six EmbrCa, compared to noncancerous testicular tissues, showed that most differentially methylated regions (DMR) in tumor tissues were methylated [39]. Of the nine genes analyzed in those DMR, eight (*AGPAT3*, *MIR1184*, *SUCLG2*, *RBMY1*, *SPANXD*, *RNF168*, *USP13* and *FAM197Y2P*) displayed reduced expression

in metastatic and nonmetastatic EmbrCa, suggesting that hypermethylation of those promoters might be associated with transcriptional repression [39]. Moreover, owing to its association with neoplastic phenotype, promoter hypermethylation of those genes might be useful for accurate identification of EmbrCa [39], although validation in a larger dataset is required and should include samples from other TGCT subtypes.

The Nodal/Cripto signaling pathway was recently shown to be involved in deregulation of germ cell homeostasis and, thus, implicated in TGCT development [40, 41]. Nodal, a TGFβ family member, signals through binding to Activin receptors in the presence of Cripto receptor [41]. Nodal signaling is transiently expressed during development of normal stem cells, playing an important role in embryogenesis for maintenance of human embryonic stem cells pluripotency, and it is absent in adult tissues [41, 42]. In TGCT, Nodal/Crypto signaling is ectopically expressed in NST, with high levels of *CRIPTO* expression generally found in EmbrCa and YST, correlating with promoter hypomethylation in EmbrCa (20%), while low levels are depicted in SE, TE and CH, correlating with promoter hypermethylation in TE and CH (72% and 75%, respectively) [40]. Because CRIPTO was detected in serum of TGCT patients it was proposed as a novel serological marker for TGCT, especially SE, EmbrCa, YST and GCNIS (sensitivity: 33, 40, 33 and 80%, respectively). Moreover, assessment of CRIPTO expression in seminal fluid is under study to enhance sensitivity of detection [40].

Ellinger *et al.* demonstrated for the first time that detection of hypermethylated cell-free serum DNA may be useful for TGCT diagnosis and prognosis [25]. Gene promoter hypermethylation in cell-free serum DNA was more frequently detected in TGCT patients than in healthy men (*APC*: 57.5 vs. 5.7%; *p16*: 53.4 vs. 17.1%; *p14*: 53.4 vs. 0%; *RASSF1A*: 46.6 vs. 0%; *PTGS2*: 45.2 vs. 0%; and *GSTP1*: 24.7 vs. 0%) [25]. Hypermethylation at any gene promoter enabled a better distinction between TGCT patients and healthy men than classical serum tumors markers, especially in patients with negativity for conventional markers (hypermethylation of two or more gene sites: 67% sensitivity and 97% specificity vs. combination of all four classical markers: 58% sensitivity) [25].

Most studies on altered DNA methylation patterns in TGCT have been performed in the class of TGCT associated with GCNIS, with only a few studies examining the methylation status in prepubertal tumors [43]. The pathogenesis of prepubertal YST differs in many aspects from that of its postpubertal counterparts [44, 45], emphasizing the need for individual analysis of each subtype. This is typified by *RUNX3* and *APC*,

which display promoter hypermethylation in prepubertal testicular YST. *RUNX3* is a tumor suppressor located at 1p36, which demonstrates frequent promoter hypermethylation (8/10 cases) and loss of heterozygosity (LOH) (6/8 cases) in prepubertal YST, but not in postpubertal YST (methylation: 0/12; LOH: 1/6) or normal young testis tissue samples (methylation: 0/10) [45]. *APC* is another tumor suppressor, mapped at 5p21, that exhibits LOH (3/9 prepubertal YST) and promoter hypermethylation (7/10 prepubertal YST vs. 0/14 prepubertal normal testis tissue) associated with loss of gene expression (in all seven prepubertal YST with *APC* promoter hypermethylation) in these rare tumors. These data suggest that both *RUNX3* and *APC* epigenetic deregulation are likely involved in prepubertal testicular YST pathogenesis [44]. The human genome is composed of less than 5% coding sequences, with repetitive sequences comprising more than 50% of the genome [46] and these are frequently hypomethylated in cancer cells [47]. The main repetitive DNA elements consist of retroelements interspersed, which correspond to about 42% of the human genome [48]. Based on their length and promoter nature, these elements can be divided into two main groups: long interspersed nucleotide elements (LINEs, 6kb in length), and small interspersed nucleotide elements (SINEs 90-300bp in size) [47]. The former comprises three main families – *LINE1*, *LINE2* and *LINE3* – of which only *LINE1* is active [46]. Among SINES, *Alu* elements are the most abundant in humans [47]. Under normal circumstances, *LINE1* and *Alu* elements are densely methylated in human tissues [48]. In TGCT *LINE1* sequences are hypomethylated in SE and NST, as well as in EmbrCa-derived cell lines, contrarily to normal peripheral blood leukocytes (PBL) and normal testicular tissue [48]. Interestingly, *Alu* repeats located at the 5' end of *CDH1* and *XIST* are hypomethylated in SE but methylated in NST, including EmbrCa cell lines, as well as in normal PBL and normal testicular tissue [48]. Moreover, demethylation of *LINE1* sequences in SE and NST, and of *Alu* repeats in SE was significantly more extensive than in the somatic cancers analyzed (testicular malignant lymphoma and renal cell carcinoma), supporting the existence of distinct regulatory patterns of demethylation of repetitive sequences, which may prove useful for diagnostic purposes [48]. The differences found for *Alu* elements methylation between SE and NST may prove clinically useful for discriminative purposes in difficult cases or follow-up of patients with mixed tumors, owing to the significant differences in prognosis and therapy regimens.

The TET proteins family (TET1-3) oxidizes 5-methylcytosine to 5-hydroximethylcytosine (5hmC) [49], which seems to be associated with epigenetic plasticity [50]. Overall, in normal hierarchically organized tissues,

the stem/progenitor cell compartment demonstrates low 5hmC levels, whereas more differentiated cells exhibit higher levels, suggesting a relationship between 5hmC and differentiation status [51]. Conversely, solid tumors are generally characterized by 5hmC reduced levels [51, 52]. Interestingly, 5hmC levels vary among different TGCT subtypes: 5hmC is absent in GCNIS, low levels are present in SE, high levels characterize TE, and intermediate levels are found in YST and EmbrCa, whereas 5hmC is detected only in stromal and Sertoli cells in the normal testis [50]. This 5hmC profile fits well with the tumorigenic model proposed for TGCT and suggests that 5hmC levels may be useful for detection and discrimination among different TGCT subtypes [50].

The vast majority of TGCT respond well to curative-intent treatment with cisplatin-based chemotherapy [29, 33], having favorable prognosis [53]. More specifically, SE tend to show excellent response to cisplatin-based chemotherapy, while NST are more variable, depending on the histological composition [29]. Nevertheless, a small proportion of patients with metastatic disease are chemoresistant, entailing disease relapse and high mortality [33, 35, 53]. Furthermore, high dose cisplatin-based chemotherapy is associated with serious cytotoxicity, including renal failure, pulmonary toxicity, vascular disease and infertility [54]. Thus, biomarkers predictive of response to chemotherapy might have important clinical usefulness. The role of DNA methylation in cisplatin-resistance (*in vitro*) has been unveiled using Tcam-2 cell line, representative of SE, exposed to 5-azacytidine [29]. Overall, demethylating treatment caused decreased resistance to cisplatin, suggesting an association between global methylation status and response to chemotherapy [29]. Therefore, therapeutic strategies combining standard chemotherapy and epidrugs may be able to overcome chemoresistance in TGCT. Another study reported on new genes capable of discriminating NST resistant to treatment with cisplatin from those that are sensitive, as well as on novel prognostic markers and therapeutic targets [35]. Interestingly, a high incidence of *RASSF1A* (52% in resistant vs. 28% in sensitive tumors) and *HIC1* (47% in resistant vs. 24% in sensitive tumors) promoter hypermethylation was associated with resistant tumors, whereas *MGMT* (31% in sensitive vs. 13% in resistant) and *RARB* (14% in sensitive vs. 0% in resistant) promoter hypermethylation was associated with tumors sensitive to cisplatin-based regimens [35]. Thus, both global and gene-specific DNA methylation analyses might constitute predictive biomarkers for TGCT.

In summary, promoter methylation status might be clinically useful for differential diagnosis of testicular masses, namely to distinguish TGCT from somatic cancers and further planning the most appropriate treatment for SE and NST. Indeed, *MGMT* and *CALCA* hypermethylation seems to correlate with NST histology (particularly EmbrCa and YST) and poor prognosis. Moreover, *CALCA* methylation association with refractory disease, unravels an opportunity for alkylating agents like temozolomide and DNA methylation inhibitors in the treatment of these tumors. Furthermore, *RUNX3* hypermethylation may be of clinical value for discriminating prepubertal from postpubertal YST, as histology alone might not distinguish those subtypes, which has implications on treatment and prognosis, as prepubertal tumors rarely metastize or recur and orchiectomy is sufficient. Finally, 5hmC levels seem to correlate well with the tumorigenic model of TGCT and might help differentiate between less (SE) and more (TE) differentiated TGCT subtypes.

Histone modifications and chromatin remodeling

Alterations in histone modifications and chromatin remodeling in TGCT have seldom been studied. There is virtually no evidence of key modifications that might be used as diagnostic or prognostic markers in this setting. Acetylation of the histone core is a process involved in DNA transcription regulation [55]. In men, this post-translational modification regulates spermatogenesis in normal and pathological testes [56]. GCNIS cells are characterized by an "open" chromatin status allowing for swift transcription, but at the expense of increased likelihood of chromosomal instability, that might foster progression into invasive TGCT. This characteristic of GCNIS is supported by low levels of DNA methylation and of repressive chromatin markers (H3K9me2, H3K27me3), as well as high levels of activating chromatin markers (H3K4me1/2/3, H3K9ac, H2A.Z), contrarily to Sertoli cells [31]. It is believed that ST originate from spermatogonia, although no histone modifications were found to match this origin [57]. Indeed, those tumors do not exhibit a preferential pattern of activating or repressive marks, making their identification exquisitely difficult [57]. A possible association between altered histone methylation patterns and deregulated expression of two genes – proto-oncogene *POU5F1* (*OCT3/4*) and tumor suppressor *RASSF1A* – has been, nonetheless, reported [58]. During normal development, *POU5F1* functions as a transcription factor involved in pluripotency regulation, being expressed in germ and embryonic stem cells [59]. In adult TGCT, this marker is specifically expressed depending on the degree of differentiation of each subtype: it is detectable in GCNIS, SE and

EmbrCa (consistent with the undifferentiated status and pluripotency) but it is absent in more differentiated tumors (YST, CH and TE) [59, 60]. As previously stated, *RASSF1A* is often silenced in TGCT due to promoter hypermethylation [33, 38] but EmbrCa cell lines NTERA2 and NCCIT exposed to a demethylating agent demonstrated only a small effect on reduction of methylation at *POU5F1* and *RASSF1A* gene promoters [58]. Nevertheless, *RASSF1A* expression was restored in association with loss of repressive mark H3K9 trimethylation and maintenance of activating mark H3K4 trimethylation in its promoter, whereas *POU5F1* expression was decreased in association with loss of activating mark H3K4 dimethylation. These findings suggest that aberrant histone modifications may be key mechanisms underlying *POU5F1* (*OCT3/4*) and *RASSF1A* deregulation in NST [58], but the clinical significance of these findings is still elusive.

Small non-coding RNAs

Altered patterns of PIWI-interacting RNAs (piRNAs) and piRNA-like transcripts have been implicated in the genesis of several cancer types, including TGCT [61]. The PIWI/piRNA pathway plays an important role in the development of male germ cells and also in TGCT tumorigenesis [62], possibly due to impaired ability to prevent chromatin instability [63]. Indeed, decreased PIWI proteins expression due to promoter hypermethylation of corresponding genes (*PIWIL1*, *PIWIL2*, *PIWIL4* and *TDRD1*), as well as of piRNAs, was previously reported in TGCT primary tissues and cell lines (SE and NST), compared to normal testis [62]. This was associated with *LINE1* hypomethylation, in agreement with the fact that *LINE1* is a target of PIWI/piRNA pathway [62]. Global piRNAs downregulation in TGCT, regardless of histologic subtype, was recently confirmed [63], and *LINE1* hypomethylation in TGCT has also been previously reported [48]. More recently, a study using samples from normal testis, TGCT, and adjacent GCNIS found that PIWI1/2/4 proteins remained expressed in GCNIS adjacent to NST, but were downregulated in GCNIS adjacent to SE [64]. As expected, *LINE1* was hypermethylated in GCNIS adjacent to NST but not in GCNIS adjacent to SE, in accordance with PIWI expression [64]. Thus, loss of *LINE1* silencing seems to be involved in the development of SE but not NST, suggesting a role for PIWI expression as diagnostic biomarker to discriminate between those tumor types [64].

Involvement of microRNAs (miRNAs) in TGCT tumorigenesis, in which they may function as oncogenes or tumor suppressors [65-67], has been previously documented, and whereas some might represent promising

therapeutic targets, others might be of diagnostic or predictive value [63]. MicroRNA-17-92 cluster has been shown to be oncogenic in GCNIS, probably due to suppression of apoptosis through inhibition of E2F1 mRNA translation, leading to progression of GCNIS to invasive TGCT [68]. Moreover, a genetic screen of primary human cells identified miR-372 and miR-373 as potential oncogenes in TGCT as well, inducing proliferation and tumorigenesis in human cells through cooperation with *RAS* and neutralization of p53-mediated CDK inhibition, due to direct targeting of tumor suppressor *LATS2* [69]. Likewise, downregulation of tumor suppressor gene *PTPN23*, due to miR-142-3p overexpression, was observed in EmbrCa-derived cell lines and SE tissues, compared to mouse spermatogonia-derived cell lines and normal testis, respectively [70]. Furthermore, high-throughput screen of 156 miRNAs in postpubertal TGCT, ST, ovarian tumors, EmbrCa-derived cell lines and normal testis identified several miRNAs that might discriminate different histological subgroups and several target genes, highlighting the importance of miR-371-373 and miR-302a-d clusters [71] and confirming previous observations [69]. Recently, miR-371-373 and miR-302 clusters overexpression was found in malignant testicular and ovarian germ cell tumors, which is consistent with downregulation of mRNAs critically involved in several biologically significant pathways [66, 72]. Those two miRNA clusters are also expressed in human testicular samples containing GCNIS and fetal gonads, further supporting that GCNIS builds up from development-arrested gonocytes that persist in the adult testis [73]. Given its strong association with TGCT, members of the miR-371-373 and miR-302-367 clusters [66, 72] may constitute good biomarkers for TGCT detection and monitoring [63], as well as possible therapeutic targets [66, 72]. Nevertheless, it should be noted that in some reports, pediatric and adult YST have been included in the same category [72], although they likely represent quite distinct diseases and this might affect the analysis of biomarker performance.

Some miRNAs deregulated in cancer and abundantly expressed in cancer tissues may be detected in body fluids owing to its stability [24], and might be useful as noninvasive serum biomarkers [26]. Pre-operative expression of miR-371-373 in serum of patients with clinical stage I TGCT was found to be significantly higher than in controls (healthy males), and decreased after orchiectomy, suggesting a role for this miRNA as a serum TGCT biomarker [24]. This is in line with the previous observation of high serum levels of members of the miR-371-373 and miR-302 clusters in a 4 year-old boy with YST, at diagnosis, and subsequent decrease during chemotherapy [74]. Moreover, significantly increased miR-367-3p, miR-371a-

3p, miR-372-3p, and miR-373-3p expression levels were found in sera from patients with TGCT compared to controls (healthy men and men with benign testicular disease) [26, 75]. In particular, serum miR-371a-3p was more sensitive (84.7%) and specific (99%) for this distinction [26], allowing for improved TGCT detection compared to classical serum markers [26, 75]. In patients with clinical stage IA disease, miR-371a-3p levels returned to basal values after orchiectomy, further supporting the tumor specificity of this miRNA [26, 76]. These results confirmed observations from other research teams [76] and the biomarker performance of serum miR-371a-3p was also emphasized due to the association between serum levels and tumor volume, absent expression in nontesticular disease and the high levels detected in testicular vein blood [26, 75-77]. Interestingly, miR-367 also appears to be a promising TGCT serum marker [26, 75], and its combination with miR-371a-3p may further increase the sensitivity and specificity of the test [77]. Remarkably, serum miR-371a-3p expression levels were recently shown to detect TGCT with diagnostic sensitivity superior to that of classical biomarkers AFP, HCG and LDH (87.8 vs. 50.4%), and an accuracy comparable to those of a panel consisting of miR-371a-3p, miR-372-3p, miR-373-3p, and miR-367-3p [78]. Importantly, this study included a large series of patients, encompassing both SE and NST (postpubertal subtypes), and it was able to replicate the findings of previous pilot studies [24, 26, 27, 75, 76, 78]. Because miR-373-3p serum levels decreased after therapy, and its persistence heralded disease persistence and/or relapse, there is great potential for patient monitoring [78]. Nevertheless, a major downside of miR-373-3p lies on its lack of expression in postpubertal TE, although the clinical impact of this finding is not clear at this point [78]. Nonetheless, since miR-373-3p expression was not tested in tumor tissue of TGCT subtypes, it is not known whether its assessment might allow for the identification of a particular TGCT subtype.

In addition to a promising role as serum biomarkers for diagnosis or monitoring, miRNAs have also been associated with TGCT sensitivity to cisplatin-based chemotherapy. Interestingly, miR-302a upregulation in cisplatin-treated EmbrCa-derived cell lines improves cisplatin-induced apoptosis through downregulation of p21 expression and p53 silencing, and increases the lethal effects of cisplatin through reduction of apoptotic threshold [54]. Hence, this synergistic effect may set the basis for a novel therapeutic approach for TGCT patients, minimizing the adverse effects of cisplatin [54]. A more comprehensive expression analysis of almost all human miRNAs in three cisplatin resistant sublines (derived from cisplatin sensitive TGCT cell lines) showed that the miR-371-373 cluster may induce resistance *in vitro*, possibly by neutralizing p53-

mediated cell death effects [79]. Many other miRNAs have been shown to be up or downregulated in those cisplatin-resistant cell lines [79], emphasizing the complexity underlying deregulated miRNAs expression in TGCT. *In vivo* analysis of these miRNAs in cisplatin-resistant tumor tissues is the next expected step that may provide useful knowledge and clinical insight on therapy-resistant TGCT.

MicroRNAs might also constitute therapeutic targets. MiR-199a displays anti-proliferative and anti-invasive effects, acting as tumor suppressor in TGCT [80] and it is downregulated, due to promoter methylation, in NT2 cell line (representative of TGCT) compared to HT cells (normal testicular fibroblasts) [81]. Remarkably, downregulation of miR-199a-3p (one of the mature forms of miR-199a) is associated with DNMT3A2 overexpression in tumor tissues (SE and EmbrCa) [82]. The therapeutic use of miR-199a-3p synthetic oligonucleotides in TGCT was suggested, because increased expression of this miRNA in NT2 cells restored *APC* and *MGMT* gene expression through promoter demethylation, possibly due to its effect on DNMT3A2 [82]. More recently, a miR-199a/miR-214/PSMD10/TP53/DNMT1 self-regulatory network was reported in TGCT, substantiating the concomitant downregulation of miR-199a along with miR-214, and *TP53* [83]. Furthermore, overexpression of DNMT1 and PSMD10 was observed in NT2 cells and EmbrCa tissues, compared to HT cells and normal tissues, respectively [83]. Thus, disruption of either pathway involving miR-199a may contribute to development of TGCT and may constitute a new therapeutic strategy [81, 83]. Other studies, addressing both testicular and ovarian germ cell tumors, have demonstrated abundant *LIN28* expression in malignant TGCT, regardless of patient age, tumor location or histological subtype, resulting in downregulation of the tumor suppressors miRNAs let-7 family, which target *MYCN* [84, 85], uncovering another potential therapeutic target.

In summary, loss of LINE1 silencing might assist clinicians in the important distinction between SE and NST; determination of miR-371-373 expression in serum both preoperatively for establishing a baseline value and post-surgery may be important for follow-up purposes; miR-371a-3p appears to be one of the most promising serological markers for TGCT detection a monitoring after treatment, with sensitivity and specificity higher than currently used markers.

Conclusion & future perspectives

Despite recent advances in our knowledge and understanding of TGCT, many questions remain unanswered: in such a treatment-responsive disease, what drives some tumors to resist and relapse? Which mechanisms underlie different behavior for neoplasms that share similar background? Furthermore, in a disease mostly afflicting young men and displaying high cure rates, issues concerning quality of life are to become more and more important. TGCT require an improved multidisciplinary and patient-centered approach for a better suited management of disease [21, 86].

Several critical questions relating to TGCT biological characteristics and clinical behavior await more definitive answers and Epigenetics may hold the key to improve current knowledge with impact on patient management through the development of novel biomarkers (Figure 2). How can epigenetic-based biomarkers improve TGCT patients' management? Biomarkers for disease detection and patient monitoring are those with more promising results thus far. Although serum DNA methylation-based markers, especially promoter methylation of *APC*, *p16*, *p14*, *RASSF1A*, *PTGS2* and *GSTP1*, have outperformed classical TGCT serum markers (AFP, HCG, LDH), the most robust candidate is undoubtedly serum miR-373-3p. Indeed, this biomarker surfaced from a series of pilot studies that coherently demonstrated its sensitivity and specificity as a TGCT biomarker, and these findings were very recently validated in a larger study from a single center. Nevertheless, external validation in a larger set of patients is required to fully confirm these promising results. As previously stated, TGCT constitute a very heterogeneous and complex group of tumors, entailing different therapeutic strategies. In this setting, accurate discrimination among TGCT subtypes might be of clinical relevance and promoter methylation of *MGMT*, *HIN1*, *RASSF1A* and *HOXA9* might be a valuable ancillary tool, eventually complemented by expression levels of *CRIPTO*, PIWI, or miR371-373 and miR302-367 clusters. It should be noted, however, that most of these results were not validated in independent datasets.

To the best of our knowledge, no epigenetic-based prognostic biomarker has been suggested for TGCT. This might be related to the overwhelmingly favorable prognosis of these neoplasms. Nevertheless, there are indeed patient subgroups with poorer prognosis, especially those resistant to cisplatin-based therapy. Thus, biomarkers predictive of that condition might be of great clinical value. Although reported findings might still be considered preliminary, promoter methylation analysis of *RASSF1A*, *HIC1*, *RARB* and *MGMT*,

as well as quantitative expression of miR-371-373 cluster show promise as predictive biomarkers in this setting. Moreover, epidrugs might also prove useful for treatment of cisplatin-resistant TGCT, in particular the demethylating agents. Finally, there are opportunities for targeting epigenetic aberrations underlying TGCT development and progression, as exemplified by miR-199a-3p synthetic oligonucleotides. It is noteworthy that prediction of treatment resistance may allow patients to avoid futile side effects and eventually benefit from alternative therapeutic modalities.

To foster the finding and validation of clinically relevant epigenetic biomarkers in TGCT it is now critical to conduct integrated analysis of possibly interconnected processes taking place in TGCT biology, such as crosslink analysis of DNA methylation, miRNAs expression, gene expression and genotype data, as recently accomplished for germ cell tumors in general [87]. In parallel, it is critical to conduct larger, multicentric, studies aimed at the validation of the most promising biomarker candidates, thriving for standardized protocols encompassing sample collection, storage and molecular analysis. It is likely that increased understanding of TGCT epigenetics will prove useful for patient management, complementing and adding information to an accurate clinical, pathological and molecular characterization of this disease [88].

miRNAs may also be used for therapeutic purposes, as exemplified by the use of miR-199a-3p synthetic oligonucleotides for targeting epigenetic aberrations in TGCT.

Executive summary

Testicular germ cell tumors

Introduction

TGCT represent over 95% of testicular neoplasms and despite accounting for only 1% of all male cancers worldwide, they comprise the most common malignancies in Caucasian men aged between 15 and 44 years-old in developed countries.

Clinical relevance

Incidence is rising, as opposed to the descendant death rates, mainly due to multimodal therapy. However, these tumors still impose several problems, namely the poor prognosis of late relapses and the emergence of treatment resistance. In such a curable disease, issues concerning quality of life, active surveillance and reproductive health are to become more relevant.

Pathological characterization of TGCT in brief

TGCT constitute a very heterogeneous group of tumors. A proper pathological evaluation is mandatory for the correct diagnosis and staging of these neoplasms. The existence of many unusual patterns makes differential diagnosis challenging.

Cancer model and recent changes in classification

There are two major categories of TGCT: those related and those unrelated to GCNIS. This classification has been recently adopted by the WHO and more closely reflects the pathogenesis of these neoplasms.

Epigenetic aberrations in TGCT and their potential as diagnostic, prognostic and predictive biomarkers

DNA methylation

- Promoter hypermethylation of *APC*, *p16*, *p14, RASSF1A, PTGS2* and *GSTP1* genes showed more utility for TGCT diagnosis and prognosis than classical serum tumors markers;
- Promoter methylation of *CALCA* is associated with refractory disease;
- There are promising results associated with gene-specific methylation, in particular for *MGMT*, *HIN1*, *RASSF1A* and *HOXA9*, which might help discriminate among tumor subtypes;
- Novel biomarkers are also being uncovered for prepubertal tumors, such as hypermethylation of *RUNX3* and *APC* genes;
- Promoter methylation status of *RASSF1A*, *HIC1*, *RARB* and *MGMT* genes suggest them as promising predictive biomarkers of TGCT response to cisplatin-based therapy**;**

Histone modifications and chromatin remodeling

- There is no evidence of major histone modifications and chromatin remodeling associated with TGCT diagnosis or prognosis. However, there is a possible relationship between altered histone methylation patterns and deregulated expression of *POU5F1 (OCT3/4)* and *RASSF1A*;

Small non-coding RNAs

- Decreased PIWI expression was suggested as a potential diagnostic biomarker for TGCT;
- The most promising and robust biomarker for diagnosis and monitoring of TGCT is serum miR-373-3p, given its high sensitivity and specificity;
- Quantitative expression of the miR-371-373 cluster might prove useful as a predictive biomarker in TGCT;
- miRNAs may also be used for therapeutic purposes, as exemplified by the use of miR-199a-3p synthetic oligonucleotides for targeting epigenetic aberrations in TGCT.

References

1. Ferlay J, Soerjomataram I, Dikshit R *et al*. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* 136(5), E359-386 (2015).
2. Oosterhuis JW, Looijenga LH. Testicular germ-cell tumours in a broader perspective. *Nat. Rev. Cancer* 5(3), 210-222 (2005).

***The paper truly reviews germ-cell tumors in a broad perspective, focusing on their origin and establishing interesting comparisons between testicular, ovarian and mediastinal germ-cell tumors.**

3. Trabert B, Chen J, Devesa SS, Bray F, Mcglynn KA. International patterns and trends in testicular cancer incidence, overall and by histologic subtype, 1973-2007. *Andrology* 3(1), 4-12 (2015).
4. Howlader N, Noone AM, Krapcho M, Miller D, Bishop K, Altekruse SF, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). SEER Cancer Statistics Review, 1975-2013, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2013/, based on November 2015 SEER data submission, posted to the SEER web site, April 2016. Accessed on April 2016.
5. Rijlaarsdam MA, Looijenga LH. An oncofetal and developmental perspective on testicular germ cell cancer. *Semin. Cancer Biol.* 29 59-74 (2014).
6. Stevenson SM, Lowrance WT. Epidemiology and Diagnosis of Testis Cancer. *Urol. Clin. North Am.* 42(3), 269-275 (2015).
7. Hanna NH, Einhorn LH. Testicular cancer--discoveries and updates. *N. Engl. J. Med.* 371(21), 2005-2016 (2014).
8. Feldman DR. Update in germ cell tumours. *Curr. Opin. Oncol.* 27(3), 177-184 (2015).

****The paper reviews the major recent advances in postpubertal testicular germ-cell tumors in a global perspective, including issues like surveillance and treatment resistance.**

9. Rajpert-De Meyts E, Mcglynn KA, Okamoto K, Jewett MA, Bokemeyer C. Testicular germ cell tumours. *Lancet* doi:10.1016/S0140-6736(15)00991-5 (2015).

***The paper reviews the major current discussions and issues about TGCT, providing an overview of their genetics, epidemiology and biology, as well as focusing on controversies and major recent improvements.**

10. Curreri SA, Fung C, Beard CJ. Secondary malignant neoplasms in testicular cancer survivors. *Urol. Oncol.* 33(9), 392-398 (2015).
11. Beyer J, Albers P, Altena R *et al*. Maintaining success, reducing treatment burden, focusing on survivorship: highlights from the third European consensus conference on diagnosis and treatment of germ-cell cancer. *Ann. Oncol.* 24(4), 878-888 (2013).
12. Kollmannsberger C, Tandstad T, Bedard PL *et al*. Patterns of relapse in patients with clinical stage I testicular cancer managed with active surveillance. *J. Clin. Oncol.* 33(1), 51-57 (2015).
13. Mir MC, Pavan N, Gonzalgo ML. Current Clinical Applications of Testicular Cancer Biomarkers. *Urol. Clin. North Am.* 43(1), 119-125 (2016).
14. Howitt BE, Berney DM. Tumors of the Testis: Morphologic Features and Molecular Alterations. *Surg. Pathol. Clin.* 8(4), 687-716 (2015).
15. Krege S, Beyer J, Souchon R *et al*. European consensus conference on diagnosis and treatment of germ cell cancer: a report of the second meeting of the European Germ Cell Cancer Consensus group (EGCCCG): part I. *Eur. Urol.* 53(3), 478-496 (2008).
16. Sharma P, Dhillon J, Agarwal G, Zargar-Shoshtari K, Sexton WJ. Disparities in Interpretation of Primary Testicular Germ Cell Tumor Pathology. *Am. J. Clin. Pathol.* 144(2), 289-294 (2015).
17. Satish K. Tickoo VER, Mahul B. Amin, Sam S. Chang, Peter A. Humphrey, James Mckiernan, John R. Srigley, Thomas M. Ulbright, Ming Zhou. Protocol for the

Examination of Specimens From Patients With Malignant Germ Cell and Sex Cord Stromal Tumors of the Testis. *College of American Pathologists* (2013).

18. Gopalan A, Dhall D, Olgac S *et al*. Testicular mixed germ cell tumors: a morphological and immunohistochemical study using stem cell markers, OCT3/4, SOX2 and GDF3, with emphasis on morphologically difficult-to-classify areas. *Mod. Pathol.* 22(8), 1066-1074 (2009).
19. Moch H, Cubilla AL, Humphrey PA, Reuter VE, Ulbright TM. The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs-Part A: Renal, Penile, and Testicular Tumours. *Eur. Urol.* doi:10.1016/j.eururo.2016.02.029 (2016).

****The paper focuses on the most recent changes in classification of TGCT, which are fundamental for understanding their genesis and biology.**

20. Holger Moch PaH, Thomas M. Ulbright, Victor E. Reuter (Eds): Who Classification of Tumours of the Urinary System and Male Genital Organs (4th Edition). Iarc: Lyon 2016.
21. Litchfield K, Levy M, Huddart RA, Shipley J, Turnbull C. The genomic landscape of testicular germ cell tumours: from susceptibility to treatment. *Nat. Rev. Urol.* doi:10.1038/nrurol.2016.107 (2016).
22. Berney DM, Looijenga L, Idrees M *et al*. Germ Cell Neoplasia In Situ (GCNIS). Evolution of the Current Nomenclature for Testicular Pre-invasive Germ Cell Malignancy. *Histopathology* doi:10.1111/his.12958 (2016).
23. Masterson TA, Rice KR, Beck SD. Current and future biologic markers for disease progression and relapse in testicular germ cell tumors: a review. *Urol. Oncol.* 32(3), 261-271 (2014).
24. Belge G, Dieckmann K-P, Spiekermann M, Balks T, Bullerdiek J. Serum levels of microRNAs miR-371-3: a novel class of serum biomarkers for testicular germ cell tumors? *Eur. Urol.* 61(5), 1068-1069 (2012).
25. Ellinger J, Albers P, Perabo FG, Müller SC, Von Ruecker A, Bastian PJ. CpG island hypermethylation of cell-free circulating serum DNA in patients with testicular cancer. *J. Urol.* 182(1), 324-329 (2009).
26. Syring I, Bartels J, Holdenrieder S, Kristiansen G, Müller SC, Ellinger J. Circulating serum miRNA (miR-367-3p, miR-371a-3p, miR-372-3p and miR-373-3p) as biomarkers in patients with testicular germ cell cancer. *J. Urol.* 193(1), 331-337 (2015).

****The paper focuses on identification of several miRNAs in serum of TGCT patients particularly miR-371a-3p which exhibited high diagnostic sensitivity and specificity .**

27. Murray MJ, Bell E, Raby KL *et al*. A pipeline to quantify serum and cerebrospinal fluid microRNAs for diagnosis and detection of relapse in paediatric malignant germ-cell tumours. *Br. J. Cancer* 114(2), 151-162 (2016).
28. Murray MJ, Coleman N. Testicular cancer: a new generation of biomarkers for malignant germ cell tumours. *Nat. Rev. Urol.* 9(6), 298-300 (2012).
29. Wermann H, Stoop H, Gillis AJ *et al*. Global DNA methylation in fetal human germ cells and germ cell tumours: association with differentiation and cisplatin resistance. *J. Pathol.* 221(4), 433-442 (2010).
30. Smiraglia DJ, Szymanska J, Kraggerud SM, Lothe RA, Peltomaki P, Plass C. Distinct epigenetic phenotypes in seminomatous and nonseminomatous testicular germ cell tumors. *Oncogene* 21(24), 3909-3916 (2002).
31. Almstrup K, Nielsen J, Mlynarska O *et al*. Carcinoma in situ testis displays permissive chromatin modifications similar to immature foetal germ cells. *Br. J. Cancer* 103(8), 1269-1276 (2010).
32. Netto GJ, Nakai Y, Nakayama M *et al*. Global DNA hypomethylation in intratubular germ cell neoplasia and seminoma, but not in nonseminomatous male germ cell tumors. *Mod. Pathol.* 21(11), 1337-1344 (2008).
33. Koul S, Houldsworth J, Mansukhani MM *et al*. Characteristic promoter hypermethylation signatures in male germ cell tumors. *Mol. Cancer* 1(1), 1 (2002).

34. Kawakami T, Okamoto K, Ogawa O, Okada Y. XISTunmethylated DNA fragments in male-derived plasma as a tumour marker for testicular cancer. *The Lancet* 363(9402), 40-42 (2004).
35. Koul S, Mckiernan JM, Narayan G *et al*. Role of promoter hypermethylation in Cisplatin treatment response of male germ cell tumors. *Mol. Cancer* 3(1), 1 (2004).
36. Smith-Sørensen B, Lind G, Skotheim R *et al*. Frequent Promoter Hypermethylation of the O 6-Methylguanine-DNA Methyltransferase (MGMT) Gene in Testicular Cancer. *J. Urol.* 175(1), 154-155 (2006).
37. Martinelli CM, Lengert AV, Carcano FM *et al*. MGMT and CALCA promoter methylation are associated with poor prognosis in testicular germ cell tumor patients. *Oncotarget* doi:10.18632/oncotarget.11167 (2016).
38. Lind G, Skotheim R, Fraga M, Abeler V, Esteller M, Lothe R. Novel epigenetically deregulated genes in testicular cancer include homeobox genes and SCGB3A1 (HIN-1). *J. Pathol.* 210(4), 441-449 (2006).
39. Cheung H-H, Yang Y, Lee T-L, Rennert O, Chan W-Y. Hypermethylation of genes in testicular embryonal carcinomas. *Br. J. Cancer* 114(2), 230-236 (2016).

****This recent paper show different patterns of expression of Cripto related with promoter methylation, depending on histological subtype of TGCT and providing a potential serological marker for TGCT diagnosis/prognosis.**

40. Spiller CM, Gillis AJ, Burnet G *et al*. Cripto: Expression, epigenetic regulation and potential diagnostic use in testicular germ cell tumors. *Mol. Oncol.* (2015).
41. Spiller CM, Bowles J, Koopman P. Nodal/Cripto signaling in fetal male germ cell development: implications for testicular germ cell tumors. *Int. J. Dev. Biol.* 57(2-4), 211-219 (2013).
42. Shen MM. Nodal signaling: developmental roles and regulation. *Development* 134(6), 1023-1034 (2007).
43. Amatruda JF, Ross JA, Christensen B *et al*. DNA methylation analysis reveals distinct methylation signatures in pediatric germ cell tumors. *BMC Cancer* 13(1), 1 (2013).
44. Kato N, Tamura G, Fukase M, Shibuya H, Motoyama T. Hypermethylation of the RUNX3 gene promoter in testicular yolk sac tumor of infants. *Am. J. Pathol.* 163(2), 387-391 (2003).
45. Kato N, Shibuya H, Fukase M, Tamura G, Motoyama T. Involvement of adenomatous polyposis coli (APC) gene in testicular yolk sac tumor of infants. *Hum. Pathol.* 37(1), 48-53 (2006).
46. Lander ES, Linton LM, Birren B *et al*. Initial sequencing and analysis of the human genome. *Nature* 409(6822), 860-921 (2001).
47. Wilson AS, Power BE, Molloy PL. DNA hypomethylation and human diseases. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer* 1775(1), 138-162 (2007).
48. Ushida H, Kawakami T, Minami K *et al*. Methylation profile of DNA repetitive elements in human testicular germ cell tumor. *Molecular carcinogenesis* 51(9), 711-722 (2012).
49. Kohli RM, Zhang Y. TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* 502(7472), 472-479 (2013).
50. Munari E, Chaux A, Vaghasia AM *et al*. Global 5-Hydroxymethylcytosine Levels Are Profoundly Reduced in Multiple Genitourinary Malignancies. *PloS one* 11(1), (2016).
51. Haffner MC, Chaux A, Meeker AK *et al*. Global 5-hydroxymethylcytosine content is significantly reduced in tissue stem/progenitor cell compartments and in human cancers. *Oncotarget* 2(8), 627-637 (2011).
52. Jin S-G, Jiang Y, Qiu R *et al*. 5-Hydroxymethylcytosine is strongly depleted in human cancers but its levels do not correlate with IDH1 mutations. *Cancer Res.* 71(24), 7360-7365 (2011).

53. Koychev D, Oechsle K, Bokemeyer C, Honecker F. Treatment of patients with relapsed and/or cisplatin-refractory metastatic germ cell tumours: an update. *Int. J. Androl.* 34(4pt2), e266-e273 (2011).
54. Liu L, Lian J, Zhang H *et al*. MicroRNA-302a sensitizes testicular embryonal carcinoma cells to cisplatin-induced cell death. *J. Cell. Physiol.* 228(12), 2294-2304 (2013).
55. Füllgrabe J, Kavanagh E, Joseph B. Histone onco-modifications. *Oncogene* 30(31), 3391-3403 (2011).
56. Faure A, Pivot-Pajot C, Kerjean A *et al*. Misregulation of histone acetylation in Sertoli cell-only syndrome and testicular cancer. *Mol. Hum. Reprod.* 9(12), 757-763 (2003).
57. Kristensen DG, Mlynarska O, Nielsen JE, Jacobsen GK, Rajpert-De Meyts E, Almstrup K. Heterogeneity of chromatin modifications in testicular spermatocytic seminoma point toward an epigenetically unstable phenotype. *Cancer Genet.* 205(9), 425-431 (2012).
58. Lambrot R, Kimmins S. Histone methylation is a critical regulator of the abnormal expression of POU5F1 and RASSF1A in testis cancer cell lines. *Int. J. Androl.* 34(2), 110-123 (2011).
59. Looijenga LH, Stoop H, De Leeuw HP *et al*. POU5F1 (OCT3/4) identifies cells with pluripotent potential in human germ cell tumors. *Cancer Res.* 63(9), 2244-2250 (2003).
60. Palumbo C, Van Roozendaal K, Gillis AJ *et al*. Expression of the PDGF α-receptor 1.5 kb transcript, OCT-4, and c-KIT in human normal and malignant tissues. Implications for the early diagnosis of testicular germ cell tumours and for our understanding of regulatory mechanisms. *J. Pathol.* 196(4), 467-477 (2002).
61. Esteller M. Non-coding RNAs in human disease. *Nat. Rev. Genet.* 12(12), 861-874 (2011).
62. Ferreira HJ, Heyn H, Garcia Del Muro X *et al*. Epigenetic loss of the PIWI/piRNA machinery in human testicular tumorigenesis. *Epigenetics* 9(1), 113-118 (2014).
63. Rounge T, Furu K, Skotheim R, Haugen T, Grotmol T, Enerly E. Profiling of the small RNA populations in human testicular germ cell tumors shows global loss of piRNAs. *Mol. Cancer* 14(1), 1 (2015).
64. Gainetdinov IV, Kondratieva SA, Skvortsova YV *et al*. Distinguishing epigenetic features of preneoplastic testis tissues adjacent to seminomas and nonseminomas. *Biomarkers* 3 4 (2016).
65. Garzon R, Calin GA, Croce CM. MicroRNAs in cancer. *Annu. Rev. Med.* 60 167-179 (2009).
66. Palmer RD, Murray MJ, Saini HK *et al*. Malignant germ cell tumors display common microRNA profiles resulting in global changes in expression of messenger RNA targets. *Cancer Res.* 70(7), 2911-2923 (2010).
67. Calin GA, Croce CM. MicroRNA signatures in human cancers. *Nat. Rev. Cancer* 6(11), 857-866 (2006).
68. Novotny GW, Sonne S, Nielsen JE *et al*. Translational repression of E2F1 mRNA in carcinoma in situ and normal testis correlates with expression of the miR-17-92 cluster. *Cell Death Differ.* (2007).
69. Voorhoeve PM, Le Sage C, Schrier M *et al*. A genetic screen implicates miRNA-372 and miRNA-373 as oncogenes in testicular germ cell tumors. *Cell* 124(6), 1169-1181 (2006).
70. Tanaka K, Kondo K, Kitajima K, Muraoka M, Nozawa A, Hara T. Tumor-suppressive Function of Protein-tyrosine Phosphatase Non-receptor Type 23 in Testicular Germ Cell Tumors Is Lost upon Overexpression of miR142–3p microRNA. *J. Biol. Chem.* 288(33), 23990-23999 (2013).
71. Gillis A, Stoop H, Hersmus R *et al*. High-throughput microRNAome analysis in human germ cell tumours. *J. Pathol.* 213(3), 319-328 (2007).
72. Murray MJ, Saini HK, Van Dongen S *et al*. The two most common histological subtypes of malignant germ cell tumour are distinguished by global microRNA profiles, associated with differential transcription factor expression. *Mol. Cancer* 9(1), 1 (2010).
73. Novotny GW, Belling KC, Bramsen JB *et al*. MicroRNA expression profiling of carcinoma in situ cells of the testis. *Endocr. Relat. Cancer* 19(3), 365-379 (2012).

74. Murray MJ, Halsall DJ, Hook CE, Williams DM, Nicholson JC, Coleman N. Identification of microRNAs from the miR-371∼ 373 and miR-302 clusters as potential serum biomarkers of malignant germ cell tumors. *Am. J. Clin. Pathol.* 135(1), 119-125 (2011).
75. Gillis AJ, Rijlaarsdam MA, Eini R *et al*. Targeted serum miRNA (TSmiR) test for diagnosis and follow-up of (testicular) germ cell cancer patients: a proof of principle. *Mol. Oncol.* 7(6), 1083-1092 (2013).
76. Dieckmann K, Spiekermann M, Balks T *et al*. MicroRNAs miR-371-3 in serum as diagnostic tools in the management of testicular germ cell tumours. *Br. J. Cancer* 107(10), 1754-1760 (2012).
77. Spiekermann M, Belge G, Winter N *et al*. MicroRNA miR-371a-3p in serum of patients with germ cell tumours: evaluations for establishing a serum biomarker. *Andrology* 3(1), 78-84 (2015).

****The paper focuses on detection of miR-371a-3p in serum of patients with TGCT, showing that it is a novel biomarker with both high sensitivity and specificity and also correlating with tumor burden.**

78. Dieckmann KP, Radtke A, Spiekermann M *et al*. Serum Levels of MicroRNA miR-371a-3p: A Sensitive and Specific New Biomarker for Germ Cell Tumours. *Eur. Urol.* doi:10.1016/j.eururo.2016.07.029 (2016).
79. Port M, Glaesener S, Ruf C *et al*. Micro-RNA expression in cisplatin resistant germ cell tumor cell lines. *Mol. Cancer* 10(1), 1 (2011).
80. Cheung H, Davis A, Lee T *et al*. Methylation of an intronic region regulates miR-199a in testicular tumor malignancy. *Oncogene* 30(31), 3404-3415 (2011).
81. Gu S, Cheung HH, Lee TL, Lu G, Poon WS, Chan WY. Molecular mechanisms of regulation and action of microRNA-199a in testicular germ cell tumor and glioblastomas. *PloS One* 8(12), e83980 (2013).
82. Chen B-F, Gu S, Suen Y-K, Li L, Chan W-Y. microRNA-199a-3p, DNMT3A, and aberrant DNA methylation in testicular cancer. *Epigenetics* 9(1), 119-128 (2014).
83. Chen B-F, Suen Y-K, Gu S, Li L, Chan W-Y. A miR-199a/miR-214 self-regulatory network via PSMD10, TP53 and DNMT1 in testicular germ cell tumor. *Sci. Rep.* 4 (2014).
84. Murray MJ, Saini HK, Siegler CA *et al*. LIN28 Expression in malignant germ cell tumors downregulates let-7 and increases oncogene levels. *Cancer Res.* 73(15), 4872-4884 (2013).
85. West JA, Viswanathan SR, Yabuuchi A *et al*. A role for Lin28 in primordial germ-cell development and germ-cell malignancy. *Nature* 460(7257), 909-913 (2009).
86. Purshouse K, Watson RA, Church DN *et al*. Value of Supraregional Multidisciplinary Review for the Contemporary Management of Testicular Tumors. *Clin. Genitourin. Cancer* doi:10.1016/j.clgc.2016.05.005 (2016).
87. Poynter JN, Bestrashniy JR, Silverstein KA *et al*. Cross platform analysis of methylation, miRNA and stem cell gene expression data in germ cell tumors highlights characteristic differences by tumor histology. *BMC Cancer* 15(1), 769 (2015).
88. Aoun F, Kourie HR, Albisinni S, Roumeguere T. Will Testicular Germ Cell Tumors Remain Untargetable? *Target. Oncol.* doi:10.1007/s11523-016-0439-4 (2016).

Figure legends

Figure 1 – Testicular germ cell tumors subtypes. A-C – Tumors unrelated to germ cell neoplasia *in situ* (A – Teratoma, prepubertal-type: epidermoid cyst; B – Yolk-sac tumor, prepubertal-type; C – Spermatocytic tumor). D-I – Tumors derived from germ cell neoplasia *in situ* (D – Teratoma, postpubertal-type; E – Yolk-sac tumor, postpubertal-type; F – Seminoma; G – Embryonal carcinoma; H – Choriocarcinoma; I – Mixed germ cell tumor: Teratoma [75%] and Seminoma [25%]). [Hematoxylin and Eosin (H&E) staining] .J – Macroscopical aspect of a Seminoma.

Figure 2 - Potential epigenetic markers for detection of testicular germ cell tumors. (Genehyper, Gene promoter hypermethylation; Genehypo, Gene promoter hypomethylation; miR, microRNA).

Table 1- Diagnostic information provided by epigenetic-based biomarkers in testicular germ cell tumors.

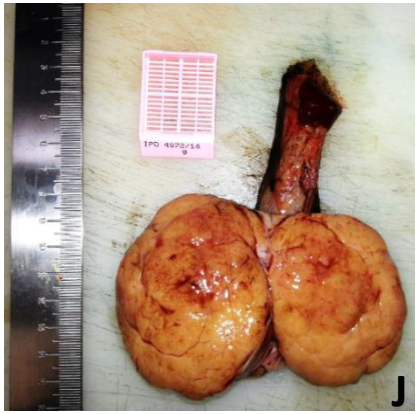
Cancer detection			
	Biomarker	Epigenetic deregulation	Sample-type details
DNA methylation	*XIST* (97% TGCT tissues (SE and NST) and 64% plasma samples) [34]	Unmethylated DNA XIST fragments (SE and NST)	Primary TGCT tissue: n=31 (18 SE and 13 NST); Plasma from 31 patients with TGCT: n=25; Controls: Peripheral blood lymphocytes from healthy males and females; Plasma from 24 male non-TGCT patients (14 renal cell carcinoma and 10 bladder carcinomas)
	MGMT (20.7% [33] and 46% [36]) *RARB* (7.6%) [33] *RASSF1A* (21.7%) [33] *MLH1* (4.3%) [33]	Promoter hypermethylation (SE and NST)	TGCT tissue: n=92 (29 SE, 44 NST and 19 mixed tumors); Controls: DNA and RNA isolated from 4 normal testes [33] Blood and primary TGCT: n=70 [33 SE (25 from patients with pure SE and 8 from patients with mixed tumors)], 35 NST and 2 GCNIS [36]
	HOXA9 (25.7% NST vs. 0% SE), *RASSF1A* (28.6% NST vs. 0% SE), *SCGB3A1/HIN-I* (54.3% NST vs. 0% SE) [38]	Promoter hypermethylation (NST)	TGCT tissue: n=61 - DNA from 55 primary TGCT: 20 SE and 35 NST (16 EmbrCa, 9 TE, 6 YST, 1 CH, 3 mixed tumors); - 7 GCNIS lesions; Cell lines: n=3; Controls: normal tissue from 4 testes
	CRIPTO (75%, 72% and 20% of methylation in CH, TE and EmbrCa, respectively) [40]	Promoter hypermethylation associated with low expression of gene (CH and TE) Promoter hypomethylation associated with high gene expression (EmbrCa)	TGCT tissue: n=35 [7 SE and 28 NST (5 CH, 4 TE, 9 EmbrCa, 10 YST)]; Controls: 10 JKT-1 cell line
	RUNX3 (80%) [44]	Promoter hypermethylation (YST in infants)	Testicular pure YST tissue: n=10; Controls: 12 NST of adulthood; 10 testes of young individuals
	APC (70%) [45]	Promoter hypermethylation (YST in infants)	Testicular pure YST tissue: n=10; Controls: 16 normal testes (2 fetal testes and 14 infantile testes)
MicroRNAs (miRNAs)	miR-371a-3p: a) 84.7% sensitivity and 99% specificity [26] b) 88.7% sensitivity and 93.4% specificity [78]	Increased in patients with TGCT compared to healthy individuals and patients with benign testicular disease	a) Serum of patients with TGCT: n=59 (40 SE and 19 NST); and with nonmalignant disease: n=17; Controls: 84 healthy male individuals [26] b) Serum of patients with TGCT: n=166 (96 SE, 60 NST and 10 TGCT of patients with relapse); Controls: 106 male individuals (12 healthy and 94 with benign testicular diseases) [78]

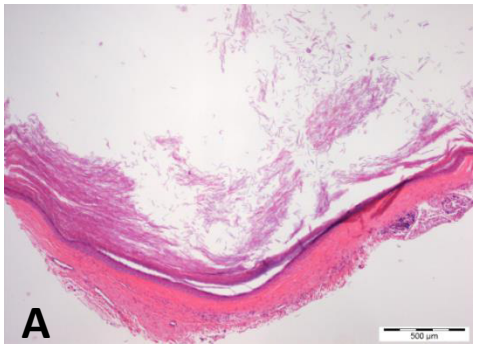
CH: Choriocarcinoma; EmbrCa: Embryonal carcinoma; GCNIS: Germ cell neoplasia *in situ*; NST: Nonseminomatous tumors; SE: Seminomas; TE: Teratoma; TGCT: Testicular germ cell tumors; YST: Yolk-sac tumor.

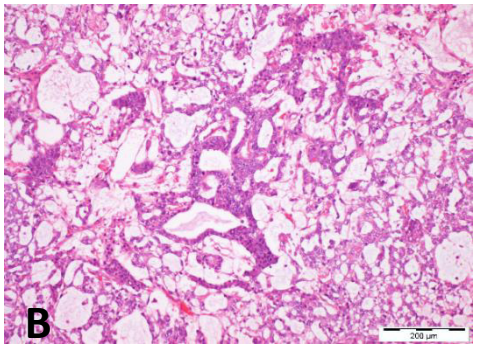
Table 2- Prognostic and predictive information provided by epigenetic-based biomarkers in testicular germ cell tumors.

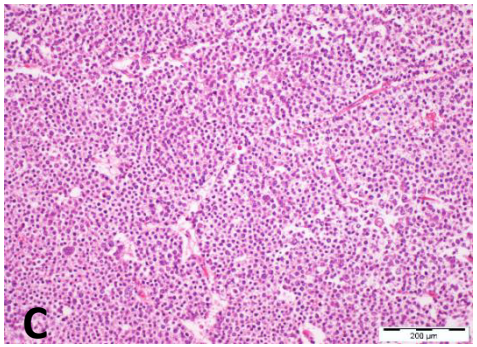
Prognostic/Predictive			
	Biomarker	Epigenetic deregulation	Sample-type details
DNA methylation	*RASSF1A* (52% in resistant vs. 28% in sensitive) [35] *HIC1* (47% in resistant vs. 24% in sensitive) TGCT [35]	Promoter hypermethylation associated with NST resistance to cisplatin-based combination chemotherapy	NST tissue: n=70 from 60 patients (31 samples of sensitive tumors obtained from 29 patients and 39 samples of resistant tumors obtained from 31 patients)
	MGMT (31% in sensitive vs. 13% in resistance) [35] *RARB* (14% in sensitive vs. 0% in resistance) [35]	Promoter hypermethylation associated with NST sensitivity to cisplatin-based combination chemotherapy	NST tissue: n=70 from 60 patients (31 samples of sensitive tumors obtained from 29 patients and 39 samples of resistant tumors obtained from 31 patients)
	MGMT (EFS = 50.5% in patients with methylated promoter vs. EFS = 77.1% in patients with unmethylated promoter) [37] *CALCA* (EFS = 51.3% in patients with methylated promoter vs. EFS = 77.0% in patients with unmethylated promoter) [37]	Promoter methylation associated with lower 5-year EFS	Primary TGCT tissues: n=72 (20 SE and 52 NST)

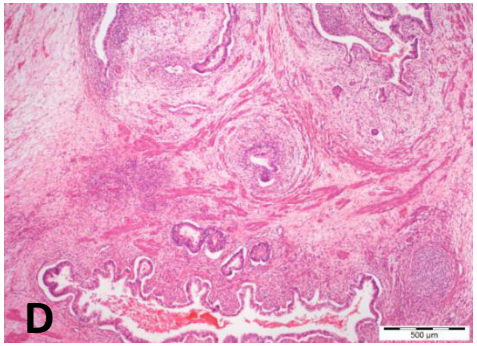
NST: Nonseminomatous tumors; TGCT: Testicular germ cell tumors; EFS: Event-free survival.

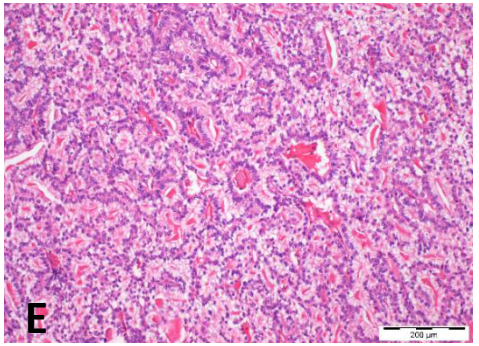

J

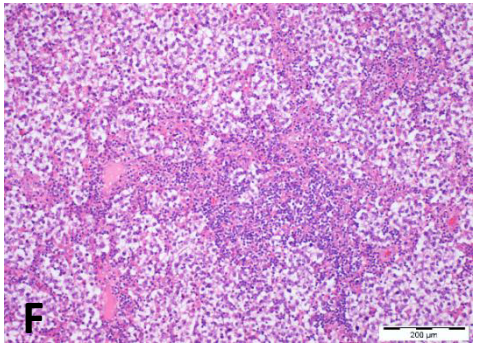

A

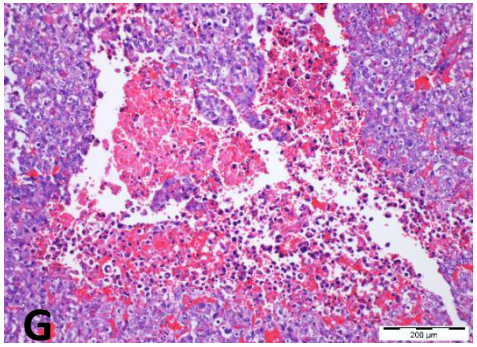

B

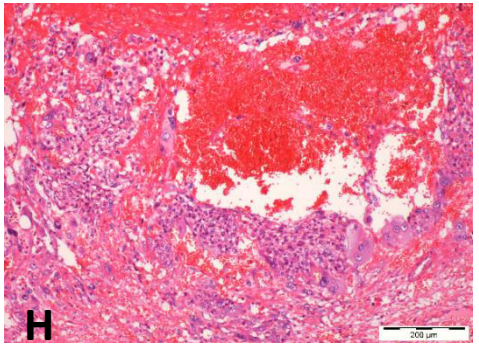

C

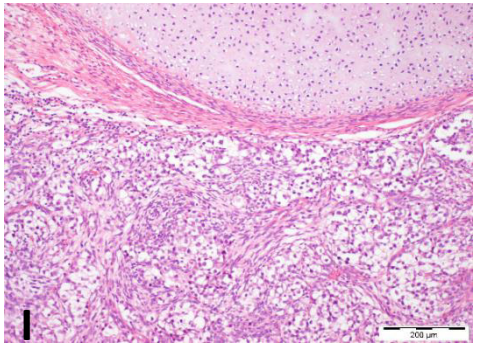

D

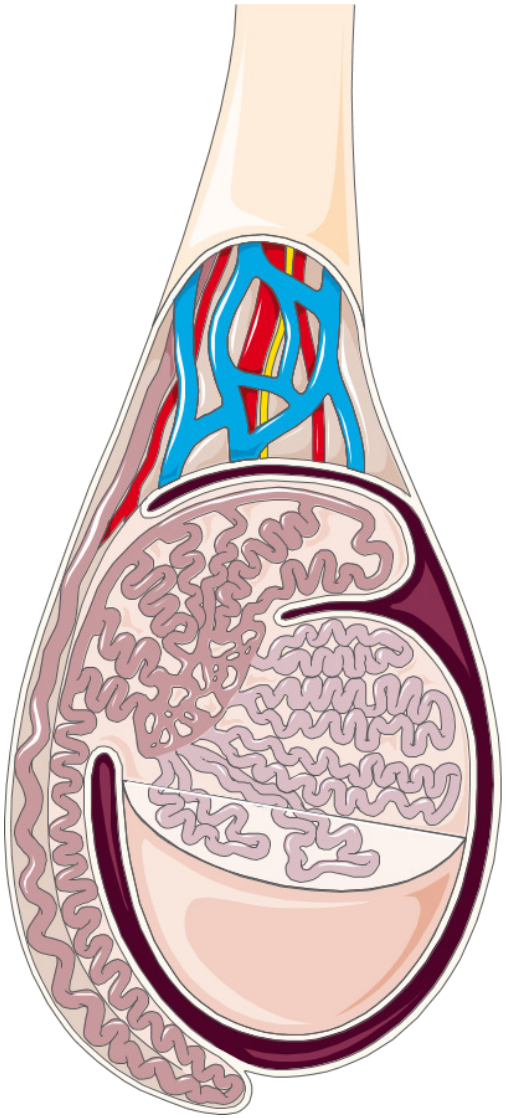

E

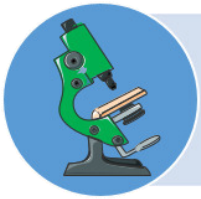

F


G


H


I





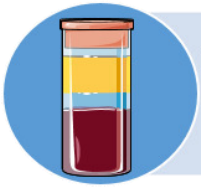
Tissue

Aberrant methylation:

XIST^{hypo}; *RASSF1A*^{hyper}; *MGMT*^{hyper}; *RARB*^{hyper}; *MLH1*^{hyper}; *SCGB3A1* (*HIN-1*)^{hyper}; *HOXA9*^{hyper}; *CRIPTO*$^{hypo/hyper}$; *LINE1*$^{hypo/hyper}$.

MicroRNAs:

miR-371-373; miR-302-367.



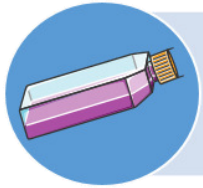
Plasma & Serum

Aberrant methylation:

XIST^{hypo}; *APC*^{hyper}; *p16*^{hyper}; *p14*^{hyper}; *RASSF1A*^{hyper}; *PTGS2*^{hyper}; *GSTP1*^{hyper}.

MicroRNAs:

miR-371a-3p; miR-367.



Cell lines

Aberrant methylation:

LINE1^{hypo}.

ANNEX II

DNA extraction from formalin-fixed paraffin-embedded tissue

DNA was extracted from FFPE tissue samples obtained from orchiectomy specimens of patients with and without TGCT, using FFPE RNA/DNA Purification Plus Kit (Norgen, Canada).

Representative block of each case was previously selected by a pathologist and tumor areas were also delimited by the same in slides stained with H&E. Sections of 10 µm were cut from FFPE block using a microtome and the tumor areas of interest were macrodissected from these sections using a disposable sterile scalpel blade and transferred to 1.5 mL safe-lock tubes. The number of macrodissected slides varied between 6 and 20, according to different tumor area corresponding to each case.

The first step of DNA extraction consists in desparaffinization of the FFPE samples through a series of xylene (VWR Chemicals, France) and absolute ethanol (Merck, Germany) washes. In this way, 1 mL of xylene (VWR Chemicals, France) was added to the sample followed by 10 minutes of incubation at 50°C and centrifugation at 14,000 x g (approximately 10,000 rpm) for 2 minutes. After carefully removing of xylene (VWR Chemicals, France) without dislodging the pellet, 1 mL of absolute ethanol (Merck, Germany) was added to the sample followed by centrifugation at 14,000 x g (approximately 10,000 rpm) for 2 minutes and its carefully removal. This step was repeated a second time and then, the pellet was left to air to dry for variable time (usually between 2 and 3 hours).

Next, the FFPE samples were digested by incubation at 55°C for 15 minutes with 300 µL of Digestion Buffer A and 10 µL of Proteinase K (20 mg/mL) (Nzytech, Portugal) followed by placing the sample on ice for 3 minutes and centrifugation at 14,000 x g (approximately 10,000 rpm) for 3 minutes. The lysate containing RNA was then collected for RNA purification following a specific protocol (not detailed here), while the pellet containing DNA was further digested for DNA by incubation the sample at 55°C overnight with 300 µL of Digestion Buffer A and 50 µL of Proteinase K (20 mg/mL) (Nzytech, Portugal), adding more Proteinase K, if necessary, until digestion was completed. After, the samples were incubated at 90°C for 2 hours followed by 3 minutes of cool down in ice. In order to remove any trace of RNA, 4 µL of RNase A (10 mg/mL) (Sigma-Aldrich, Germany) was added to the cooled lysate followed by 5 minutes of incubation at room temperature. 300 µL of Buffer RL and 250 µL absolute ethanol (Merck, Germany) were added to the lysate and the solution was loaded onto a DNA Purification Micro Colum in two steps, where in each one of them up to 600 µL of lysate were applied onto the column and centrifuged at 14,000 x g

(approximately 10,000 rpm) for 1 minute. Then, 3 washes steps with 600 µL of Wash Solution A were performed. In each of these steps, after added the 600 µL of Wash Solution A, the column was centrifuged at 14,000 x g (approximately 10,000 rpm) for 1 minute and the flowthrough was discarded. Lastly, the purified DNA was eluted in 20 µL of the provided Elution Buffer F [divided in two steps of 10 µL of this buffer followed by centrifugation at 14,000 x g (approximately 10,000 rpm) for 1 minute].

After elution, for each sample, the DNA concentrations and respective purifications ratios were measured by NanoDrop Lite Spectrophotometer (NanoDrop Technologies, USA) and stored at -20°C.

ANNEX III

Sodium bisulfite treatment of DNA

Sodium bisulfite treatment of extracted and quantified DNA was performed by EZ DNA Methylation-Gold™ Kit (Zymo Research, CA, USA). The DNA volume used in this process was initially calculated according to its concentration, wherein the quantity of bisulfite-modified genomic DNA varied between 250 and 1200 nanograms. According to different DNA volume used for each sample, we added sterile bidistilled water (B.Braun, Melsungen, Germany) up to a total volume of 20 µL. For each sample, was also added 130 µL of CT conversion reagent solution followed by incubation in Veriti® Thermal Cycler (Applied Biosystems, CA, USA) for 10 minutes at 98°C and 3 hours at 64°C (3 cycles of 60 minutes each), allowing denaturation and sodium bisulfite conversion of DNA. After incubation, each sample was transferred to corresponding Zymo-Spin IC™ column with 600 µL of M-binding buffer and incubated at room temperature for 10 minutes, followed by centrifugation at 10,000 rpm for 30 seconds. After that, one wash step was performed by adding 100 µL of M-Wash buffer and centrifuged at 10,000 rpm for 30 seconds. The next step consists in desulphonation by incubation with 200 µL M-Desulphonation buffer at room temperature for 20 minutes, followed by a centrifugation at 10,000 rpm for 30 seconds. Next, two washes steps were achieved adding 200 µL of M-Wash buffer followed by centrifugation at 10,000 rpm for 30 seconds in each of these steps. Lastly, the columns were placed in a 1.5 mL safe-lock tube and DNA was eluted by 5 minutes of incubation at room temperature with a volume of sterile bidistilled water (B.Braun, Melsungen, Germany) variable between 6.25 and 30 µL, according to different quantity of intended modified-DNA, followed by a centrifugation at 12,000 rpm for 30 seconds. This step was repeated a second time to obtain a total volume of bisulfite-modified genomic DNA ranging between 12.5 and 60 µL, which was then stored at -80°C.

CpGenome™ Universal Methylated DNA (Merck Millipore, Germany) was also modified using this process. In the last step, this DNA was eluted in a total volume of 20 µL of sterile bidistilled water (B.Braun, Melsungen, Germany).

ANNEX IV

Table 18. Serum tumor markers positivity in GCNIS-related TGCT.

Elevated serum tumor markers						
TGCT subtypes	**AFP**		**β-hCG**		**LDH**	
	Pre-orchiectomy	Post-orchiectomy	Pre-orchiectomy	Post-orchiectomy	Pre-orchiectomy	Post-orchiectomy
GCNIS-related	49/140	31/124	72/139	29/124	48/112	27/105

TGCT: Testicular germ cell tumors; GCNIS: Germ cell neoplasia *in situ*; AFP: Alpha-fetoprotein; β-hCG: Beta-human chorionic gonadotropin; LDH: Lactate dehydrogenase.